MINUTES OF THE CENSUS ADVISORY COMMITTEE
OF THE PROFESSIONAL ASSOCIATIONS

(This Committee consists of members of the American Economic Association (AEA), the American Marketing Association (AMA), the American Statistical Association (ASA), and the Population Association of America (PAA).)

At the Embassy Suites, Washington, DC, April 23-24, 1998

CONTENTS

Page

Minutes Prepared by the History Staff

ATTENDANTS AT MEETING

(Asterisk [*] indicates one day only)

Members Present

Frederic M. Scherer, Chairperson

Tony Adams, AMA
Laurie Ashcraft, AMA
Patricia Becker, PAA
Robert Bell, ASA
Roger R. Betancourt, AEA
Michael Etzel, AMA
Joseph Garrett, ASA*
Malay Ghosh, ASA
Michael Gort, AEA
Linda Jacobsen, PAA
Katherine Jocz, AMA*
Jacob Klerman, PAA
Lee Lillard, AEA
Rebecca A. Maynard, AEA
Dowell Myers, PAA
William O'Hare, ASA*
Rosann Spiro, AMA*
Elizabeth A. Stasny, ASA
Lynne Stokes, ASA
Ross Stolzenberg, PAA*
Roger E. Tourangeau, ASA
Paul R. Voss, PAA
Robert Willis, AEA
Franklin D. Wilson, PAA

Members Absent

Ernst R. Berndt, AEA
David A. Binder, ASA
Lynn E. Browne, AEA
Beth Fischer, AMA
F. Thomas Juster, ASA
Ariel Pakes, AEA
Arthur Redmond, AMA
Debra Semans, AMA
David Stewart, AMA
James Trussell, PAA

Other Persons Present

Michelle Ash, Minority Counsel, U.S. House of Representatives
Alden Bean, Lehigh University, Center for Innovation Management Studies
Gregory A. Bischalk, Senior Economist, Appalachian Regional Commission
James H. Burow, Assistant Director, U.S. General Accounting Office
Lara Chamberlin, U.S. House of Representatives
Mickey Clayton, Management Analyst, U.S. Department of Commerce
Wesley Cohen, Professor of Economics, Carnegie Mellon University
Edwin Dean, Associate Commissioner, Bureau of Labor Statistics
Richard W. Dodge, Editor Senior Analyst, Bernan Associates
Candace Feit, Data Analyst, Bernan Associates
Linda Engelmeier, Management Analyst, U.S. Department of Commerce
David Flaherty, U.S. House of Representatives
Jeanne Griffith, National Science Foundation
Margaret Grucza, Director Research Services, Industrial Research Institute
George E. Hall, Senior Managing Editor, Bernan Associates
Peter H. Henderson, Study Director, National Research Council Office of Scientific and Engineering Personnel
Thomas Hofeller, U.S. House of Representatives
Paul Hsen, Bureau of Labor Statistics
John E. Jankowski, Director, Research and Development Program, National Science Foundation
Ron Jarmin, Center for Economic Studies
Nancy Kirkendall, Mathematical Statistician, Office of Management and Budget
Steven Landefeld, Director, Bureau of Economic Analysis
William F. Long, President, Business Performance Research Associates, Inc.
Tim Maney, U.S. House of Representatives
David McMillen, U.S. House of Representatives
Stephen A. Merrill, Executive Director, National Research Council
Dereck Orr, U.S. House of Representatives
Lawrence Rausch, National Science Foundation
Robert Shapiro, Under Secretary for Economic Affairs, U.S. Department of Commerce
Sybil Stershic, Chairman of the Board, American Marketing Association
Ralph Wyndrum, Vice President for Technology of AT&T
Raymond M. Wolfe, National Science Foundation
Paula Young, Bureau of Economic Analysis

Summary

Introductory remarks. Mr. Holmes reviewed senior staff changes at the Census Bureau and noted that Mr. Robert Shapiro has been sworn in as Under Secretary for Economic Affairs at the Department of Commerce. He announced that the Committee will take part in a meeting on race tabulations with the Census Bureau's Race and Ethnic Populations Advisory Committees, and the Department of Commerce's 2000 Census Advisory Committee on June 3, 1998. The next regular meeting of the Census Advisory Committee of the Professional Associations is scheduled for October 22-23, 1998.

In reviewing current activities of the agency, Mr. Holmes pointed out that Census 2000 State Action Committee meetings will begin in the fall of this year. Two lawsuits have been filed challenging the use of statistical sampling in Census 2000, and several parties have filed motions to intervene on behalf of the Census Bureau. With regard to the budget, the President signed the fiscal year (FY) 1998 appropriation bill for the Bureau on October 26, 1997, which included $137 million for salary and expenses and $556 million for periodic censuses and programs. The first hearings for the FY 1999 budget appropriations bill were held on March 25, 1998.

Census Bureau responses to Committee recommendations/report on October 1997 meeting. In discussion, members pointed out that the American Statistical Association (ASA) and Population Association of America (PAA) subgroups had submitted conflicting recommendations at the meeting, but had reworded them as recommendation 19 to eliminate the conflicts. The Bureau's response, however, seems to respond to the old versions of the recommendation. The agency's response to recommendation 16 is in error.

Responding to questions by members, Ms. Schneider said that while hearings have been held on data sharing among statistical agencies, no formal legislative actions have been taken. Dr. Knickerbocker added that Senate Bill S. 1404 has been introduced to create a commission to investigate consolidating parts of the Federal statistical community—specifically the Census Bureau, the Bureau of Economic Affairs (BEA), and the Bureau of Labor Statistics (BLS). The bill also expresses Congress' sentiment in support of data sharing. In reply to further questions, Dr. Knickerbocker pointed out that the greatest cost savings would be realized through the consolidation of the Census Bureau and BLS—particularly through sharing information on industrial classification.

Members noted that a published article on the subject has questioned whether data collected by the BLS in its Employment Survey 202 belongs to the individual states, or to the BLS. Some states are concerned about data sharing because their local laws forbid sharing certain data.

The Census Bureau's plans for poverty measurement. (AEA, ASA, PAA) Discussants supported the plans to use information from the Survey of Income and Program Participation (SIPP) for analyzing poverty issues. Members expressed concerns about the problem of survey panel attrition, the exclusion of medical expenditures and medical insurance premiums from income measurement, the possibility of biased estimates, and insuring that the debate over poverty measurement does not damage the Bureau's other core activities. They urged the agency to follow up survey contact refusals to improve response, study the burden of

prescription drugs on the elderly, conduct a "Consumer Expenditures Survey (CES)"-type supplement to the SIPP, and take population mobility and geographic variations into consideration in its poverty measurement. Members suggested the Bureau separate and measure the demographic and socioeconomic changes associated with welfare reform, particularly for people who move off welfare. Members expressed interest in the Bureau's plans for interfacing experimental measurements with other data series, including the decennial census, the American Community Survey (ACS), and the Small Area Income and Poverty Estimates program.

Responding to questions by members, Dr. Weinberg said it will be difficult to supplement the SIPP with a CES supplement, since respondents consider those two surveys to be among the most burdensome. The agency will use statistical matching to supplement the SIPP. The likely time frame for the new SIPP data on poverty is sometime in 2003. The final decision on the poverty threshold and income measurement will be made by the Office of Management and Budget (OMB).

Are we on the right track with the corporate marketing program? (AMA) Mr. Selby summarized key findings of the Bureau's Corporate Marketing Plan focus groups. There was an extended discussion of the Bureau's marketing plans and activities. Members commended the agency's corporate identity and logo, emphasized the importance of targeting and segmenting markets for promotion of cooperation with the Bureau's data collection and of sales of data products, and stressed the importance of developing means to make products attractive to potential customers over time to maintain a revenue flow. The agency needed to have more information about its potential audiences. Discussants did not feel the Bureau's marketing survey's primary message ("being the official source of government statistics...") was exciting, nor did it convey the idea of any benefit of cooperating with Bureau censuses or surveys, or of obtaining agency products. Members urged the Bureau to promote itself to the public as an authority on demographic information, using a variety of marketing venues. In discussions, members pointed out that data users wanted up-to-date statistics—they did not want 1990 census data in 1998—and criticized the Bureau's Current Population Survey (CPS) data as poorly organized, with outdated and inaccurate data, and uninteresting tabular presentation. The agency needs to expand user access to online data files, develop more user-friendly products, and improve technical support for its products.

Economic census update. (AEA) Mr. Govoni reviewed the status of the 1997 Economic Census. Following the census mailout in December 1997, processing began by the end of February 1998. Response is running about 3 percent short of the 1992 pace, and the agency projects a final unit response rate of 80-85 percent. During processing, the data are keyed to an electronic data file and a series of computerized checks are done to identify problems for correction and to impute data where needed. After editing, correction, and imputation, the data are tabulated and the aggregate totals reviewed before the first reports are issued. The Bureau plans to release an advance national-level report early in 1999, to be followed by approximately 500 industry and state reports.

Dr. Dunne outlined the analysis done to identify economic census response patterns. He reported that the smallest establishments (1-10 employees) had a 20-percent lower response rate than the largest (250 or more employees) establishments, that new establishments were 12 percent less likely to respond than older ones, and that response from small new companies is only about 60 percent.

In discussion, members pointed out that researchers have found that, for questions such as expenditures, supplying item detail is sometimes easier than providing a "global" figure. Responding to questions by members, Mr. Govoni said the Bureau uses industry averages from the Annual Survey of Manufactures, previous censuses, or reported averages from specific industries for its imputation work. The agency tries to identify imputed data in the records, not only for users, but for the analysts' work as well.

How can the Census Bureau get consistent and useful feedback from its customers? (AMA) In discussions, members suggested user conferences were a useful method of obtaining feedback from customers, pointing out that such conferences facilitated network building, face-to-face contact with users, and instructive workshops to provide users with details on data uses. In reply to questions by members, Mr. Wynegar said that staff time and value to the agency, rather than cost, were the issues for the Bureau when considering implementing user conferences.

Responding to the questions raised in the paper, discussants said the Bureau's concentration on electronic media for communications exchange is appropriate, but that the agency may want to forego interactive online chat rooms in favor of a delayed form of response that allows more control. The Bureau also should consider introducing a closed bulletin board coupled with a subscriber mailing list. Online feedback from customers should be kept in perspective and viewed in the context of customer-satisfaction surveys. The agency should also keep track of calls to its customer service office and summarize these calls for management review.

Overview of the Census 2000 and dress rehearsal plans. (ASA, PAA) Mr. Thompson reviewed the budget agreement concluded by the Administration and the Congress, the resulting changes to the dress rehearsal plans, and the current status of operations at the three dress rehearsal sites. He noted that the Menominee, WI, and Sacramento, CA, sites are on schedule with respect to hiring staff, while the Columbia, SC, site is behind schedule due in part to very low unemployment in that area. Promotion and outreach activities are underway, as is the evaluation of the effectiveness of those efforts. Mail enumeration is in progress, with response rates of 34.9 percent in Menominee County, WI; 44.9 percent in Sacramento, CA; and 44.5 percent in Columbia, SC. Data capture began full production at the Bureau's Jeffersonville, IN, facility on April 20. The Bureau has received nearly 16,000 telephone calls for assistance in completing the questionnaires, about half of which have been resolved using the interactive voice recognition system without referral to an operator. The Bureau completed the first phase (independent address listing) of the integrated coverage measurement program (ICM) and is matching the housing unit address list to the dress rehearsal address list. Nonresponse follow-up will begin May 14.

Turning to Census 2000, Mr. Thompson noted that the local update of census addresses is underway and about 9,000 local governments have so far agreed to participate. The Census Bureau submitted the actual questions to be used on the Census 2000 questionnaire to Congress on April 1, 1998.

Sampling and estimation in Census 2000 and the dress rehearsal. (ASA, PAA) Dr. Hogan summarized the procedures—nonresponse followup, undeliverable as addressed vacant follow-up, Integrated Coverage Measurement (ICM), and service-based

enumeration—employed at the two dress rehearsal sites at which sampling and estimation was used (Menominee County, WI, and Sacramento, CA). He noted that, in response to Committee recommendations, the Bureau had increased the sample percentage from 10 to 30 percent. For Census 2000, the Bureau's sampling plan calls for a 750,000 housing-unit sample for the ICM. The ICM will not be combined with demographic analysis estimates in Census 2000, largely because the Bureau could not identify one methodology that was clearly superior to any of the various alternatives.

In comments, members emphasized that there is almost universal agreement among demographers and statisticians that it is possible to design sampling procedures that will perform better than the procedures used in the 1990 census. The Congress and the courts will decide what the final numbers will be, but the members urged the Bureau to make every effort to preserve ICM and the Post-Enumeration Survey (PES). Members expressed concern about the feasibility of the Bureau's plans for Census 2000, particularly (1) the effort to hire large numbers of temporary workers, (2) the complexity of the estimation procedures, (3) the need for a more explicit schedule for planning, and (4) plans to use controlled rounding, and (5) setting of "drop dead" dates for mail returns.

Responding to questions by participants, Dr.Hogan said that Census 2000 data products will probably look very much like 1990 products, and that geographic prestratification is being considered for Census 2000.

Overview of indicators of innovation and technology. (AEA) Mr. Shapiro said that the most important objective of the Department and the Bureau is the development and application of professional protocols and measures for the census; the census design should be dictated by the state of the art of statistical design and method. Responding to questions by the members, he noted that the Administration agreed to expedited review of the constitutionality of using statistical sampling in the census, and is confident that the courts will uphold that position.

Dr. Cooper reviewed the National Science Foundation's workshop (held in February 1997) to improve statistical information on industrial innovation. The workshop looked at the demand for information on innovation and specific areas of policy needs. Participants made a series of recommendations to improve existing data on innovation (e.g., by linking data sets), expanding coverage of research and development (R&D), developing new research and development data, clarifying the role of the Federal and private agencies in collecting the information, and forming a consultative body to advise the various Federal agencies involved on these issues.

Dr. Haltiwanger commented that the Census Bureau was asking members to consider what the agency should do about existing R&D surveys, whether it is possible to satisfy some of the data needs with modules that could be added to existing surveys, and whether a new technology or innovation survey is required.

Responding to questions from members, Dr. Knickerbocker said a fourth alternative is to find new ways to link existing data sets on R&D to improve the data.

Panel discussion: The National Science Foundation research and development survey. (AEA) Ms. Champion described the National Science Foundation's (NSF's) research and development (R&D) survey, conducted annually by the Census Bureau. The survey collects key data variables from a sample of several thousand companies and was redesigned in 1992 to improve coverage of smaller companies' research efforts and to include nonmanufacturing industries.

In discussion, members of the panel suggested that state and geographic data are needed on R&D, and that the Bureau should analyze the effect of constant panel change for small firms, the introduction of small firms into the sample, and the introduction of high-technology firms. Tracking initial public offerings and patent data could also provide information on small high-technology companies. Discussants suggested that product breakdowns are needed, and warned that the scientific community has not agreed upon the distinction between basic and applied research.

In a discussion of using microdata files to look at R&D, Mr. Jarmon pointed out that most of the data are collected at the establishment level, which makes it difficult to allocate R&D expenditures. The Bureau/NSF survey is voluntary, so coverage of some items is "spotty," and the data are not collected with microdata research in mind.

Responding to questions by members, Mr. Jankowski said response was about 50 percent [for applied R&D by product class], and Ms. Champion pointed out that no follow-up survey was done following revision of the R&D survey's instructions.

Demonstration of the latest DADS prototype. (AMA) Ms. Rowland outlined the background of the Bureau's development program for the Data Access and Dissemination System (DADS), noting that International Business Machines (IBM) and its subcontractor Oracle and Environmental Systems Research, Inc., have been awarded the contract to build the DADS for Census 2000.

Ms. Moore demonstrated the DADS97 prototype and its different tools. She pointed out that no "stop" or "back" button is available in the prototype—once an enquiry was started, it had to be completed. The first release of the production system—DADS98—is scheduled for January 1999; DADS98 will have access to the 1990 census, the 1997 Economic Census, and the 1996 and 1997 American Community Survey (ACS) data products. The system will be upgraded with access to the Census 2000 Dress Rehearsal products in March 1999.

In response to questions by members, Ms. Moore said that users will be able to download the DADS software from the Internet, although there might be a fee involved. Mr. Wynegar pointed out that IBM will not be involved in selling the DADS software, as the system is a delivery mechanism, not a product for sale.

How do we evaluate the dress rehearsal and Census 2000? (ASA, PAA) Ms. Killion described the objectives of the Census 2000 dress rehearsal and the Bureau's plans for evaluating the operation. The agency has established eight quality review boards for the evaluation program, and will prepare and release a series of evaluation reports covering major facets of the dress rehearsal (e.g., questionnaire, the master address file, coverage measurement and evaluation, etc.).

There was an extended discussion of the Bureau's plans. Members pointed out that the statistical community supports using statistical sampling in Census 2000 and the Integrated Coverage Measurement (ICM) program for the census, and noted that the master address file (MAF) and local update of census addresses are not being tested in the dress rehearsal. Participants suggested that during Census 2000 the Bureau collect more data on different types of "last resort" information to compare data obtained in the ICM and non-ICM blocks. They added that the Bureau should have more information about prospective enumerators (e.g., work history, performance on screening tests, etc.) that could be linked to performance measures. A 30-percent increase in public awareness of the census may be too modest. Discussants pointed out that the statistical community expects that sampling for nonresponse in the census will improve timeliness and reduce cost, and that the ICM will reduce the differential undercount.

Responding to questions by members, Ms. Killion said some proposed evaluations of the dress rehearsal have been eliminated, but that the program includes evaluations of MAF coverage and MAF-building processes. Ms. Schneider noted that the Bureau may have to reconsider widespread distribution of blank questionnaires to reduce the number of completed duplicate forms.

Panel discussion: Where do we go from here? (AEA) Dr. Knickerbocker moderated a panel discussion that included presentations by panelists on (1) the criteria used by major corporations for budgeting research and development (R&D), (2) how R&D data can benefit public policy, aid decision making, and improve understanding of innovation, (3) the National Science Foundation's means of monitoring research and innovation, (4) the Bureau of Economic Analysis's decision not to fund satellite accounts (including R&D accounts), (5) the Bureau of Labor Statistics' study of the direct effects on multi-factor productivity growth, and (6) the Industrial Research Institute/Census for Innovation Management Studies' annual R&D survey.

In an extended discussion, members noted that collecting data on R&D is complicated by some companies' policy of only responding to mandatory surveys, and that so few companies respond to some categories of questions (e.g., process data) that some surveys do not attempt to collect the information. There was general agreement that line-of-business data are important and that mandatory reporting may be the only way to collect these data, while patent data were problematic, since they were easy to misinterpret. International comparability of the data was becoming more important, and the knowledge level of the respondents was critical to survey data quality. Participants expressed concern about the lack of data on outputs related to R&D, and that the R&D expenditures data available often show small companies with no budgets for research, which cannot be correct. There is a need for data on outputs as well as inputs, although the relationship between the two variables is very "noisy." Linking the data is costly and time consuming, and will require input data at the business segment level.

How do we evaluate the marketing strategy for the dress rehearsal and Census 2000? (AMA, ASA, PAA) Ms. Bates outlined the Bureau's plans for evaluating the dress rehearsal marketing operation, noting that the agency is using a two-wave survey research design administered by a random-digit dial survey.

Members pointed out that it will be difficult to determine if an increase in response and awareness resulted from the advertising campaign or the questionnaire mailout. Biases could

occur in the data due to the restricted size of the sample. The Bureau needs to identify the public attitudes that had to be changed, particularly concerns about confidentiality. Discussants expressed concern about the probable cost of the advertising strategy when applied to the census and the need to have measurements at key points of the marketing hierarchy of effects model. They suggested that a three-wave design would be needed to measure awareness independent of the effects of the mailout.

Responding to questions from members, Ms. Bates pointed out that the Bureau preferred the three-panel design, but timing and other factors prevented its adoption. The questionnaire for the second wave included additional questions intended to measure innovative and aggressive advertising at the South Carolina site. The questionnaires for the evaluation included items addressing confidentiality, and she noted that the early data from the survey indicate rather poor response rates. Dr. Meyer said the Bureau plans to change the advertising campaign for Census 2000 to include several new elements aimed at reaching targeted groups in the population.

How should the Census Bureau price data products through DADS? (AMA, PAA) Mr. Kavaliunas outlined the Bureau's review of its product pricing policies, noting that the Bureau is prohibited from entering into any restrictive or exclusive arrangements for its products, and requires that the agency recover the cost of disseminating data products.

Responding to questions by members, Mr. Kavaliunas said that some of the agency's products will be available electronically through the Data Access and Dissemination System (DADS), while others will be available in other media. Data users will be able to view some tables via the DADS on the Internet without charge. The Bureau is not permitted to charge royalties to companies using census data. The cost that could be passed on to users in sales of Census 2000 products is about $52 million.

In an extended discussion, members asserted that the Bureau's products are too expensive, suggesting that the agency should be concerned about what customers are paying for the same data from alternative sources. Current costs of Bureau products make it feasible for a company to purchase a product on CD-ROM, for example, and reproduce the product for sale at a much lower price under its own commercial name. The Bureau needs to broaden its support base and guarantee broad access to its data products while keeping costs within its appropriation for data dissemination, and needs to better understand product dissemination. Some members questioned whether the agency should be disseminating data at all, suggesting that private companies could do the job better. The budget currently spent on dissemination should go toward better data collection.

General edit and imputation research (ASA) Dr. Winkler described the Bureau's work on developing two generalized processing systems—the Structured Programs for Economic Editing and Referrals (SPEER) and the DISCRETE systems. These systems are based on the Fellegi-Holt model of editing. The agency was asking members (1) whether it should continue to develop these systems, (2) for suggestions about the direction edit/imputation research should take, and (3) for suggestions on developing additional edit modules, better user interfaces, and creating documentation and training methods.

In discussion, members suggested that the improved imputation for race and age in tests using the new systems encouraged continued research. The agency should consider

allowing for using different statistical models in different parts of the country, and needs to develop the best possible user interfaces and documentation. Discussants pointed out that there are reasons for concern about the assumptions used in the various statistical models; trying to impute only one variable at a time, then using that variable as part of the input for the next imputation, may simplify things.

In response to members comments, Dr. Winkler said the Census Bureau looked at many closely related models to identify those that gave the best predictive capability for imputation. The agency plans to run a generalized system in parallel to the more conventional production system in the dress rehearsal. He noted that the Bureau has not conducted a retrieval interview program to check the accuracy of its imputed data, but added that there is strong evidence that careful modeling will improve estimates.

Chief economist updates (AEA) Dr. Haltiwanger updated the members on the recent activities of the Center for Economic Studies (CES), noting that the planned expansion of the research data center (RDC) program is continuing, and that the CES is interested in determining what new kinds of databases can be created. He outlined the resources needed for each center, and noted that the National Science Foundation (NSF) has offered some financial support for the new RDCs. With regard to new databases, he said the CES is trying to link establishment economic data to demographic data files.

In discussion, members suggested the Bureau must think through its rules regarding access before imposing more restrictions on researchers, and pointed out that most discussions of confidentiality center on the technical issues involved, rather than the motivation of someone to violate the requirements. The Bureau should require that when data sets are created from agency files by researchers at the RDCs, those sets must meet public-use data filed standards, and must be retained by the RDCs.

Responding to comments and questions by members, Dr. Haltiwanger said that (1) no researcher at an RDC has yet violated Bureau confidentiality or other requirements, (2) data files created at the RDCs must meet public-use standards and are retained by the RDC, (3) CES plans to deliver to major data files to the RDCs this year (1982-1992 economic census files with longitudinal linkages, and the Standard Statistical Establishment List file), (4) an intermediate file with critical data fields extracted from the raw data file could be compiled for some of the Bureau's programs, and (5) agreed that creating a combined public-use data and economic data file would be a good idea.

How should we promote confidentiality in the decennial census? (AMA) Census Bureau staff made presentations on three alternative approaches to promoting confidentiality—using outside spokespersons in place of paid advertising, combining paid advertising with local partnerships in a centralized campaign, and reviewing the promotion campaigns used at the dress rehearsal sites. In discussion, members emphasized that confidentiality is critical in improving response to the census (although members also pointed out that the "what's in it for me" message also is important) and that the Bureau needs more information about the effectiveness of paid advertising to promote public trust in the confidentiality of the census data. The lack of a control cell in the dress rehearsal prevented researchers from isolating the effects of advertising. Several participants suggested that the extensive use of focus groups in the agency's review of the advertising effort was a mistake; focus groups can be used to develop and refine methods, but not to test them. Commercials

should be pretested before being broadcast. The Bureau should target specific audiences that are particularly concerned about confidentiality for a sustained advertising campaign. At least one member expressed doubt that the $100 million budget for paid advertising was sufficient for these tasks.

How will the OMB proposal on tabulation of race and ethnicity data be implemented in dress rehearsal tabulations? (ASA, PAA) Dr. del Pinal summarized the major changes made by the Office of Management and the Budget (OMB) in race and ethnic classifications used by the Federal Government. He noted that the Bureau plans for Census 2000 to release fewer products than in the 1990 census, and use the new Data Access and Dissemination System (DADS) to enable users to specify the tabulations they want for different racial and ethnic categories, geographic units, subject matter, etc. Dr. Tucker reviewed OMB's efforts to development new guidelines for race and ethnic classification. The agency has formed two working groups to consider the policy needs of Federal agencies and the technical questions of data tabulation, and expect to release preliminary reports in June 1998.

In comments, members supported efforts to improve race reporting, but expressed concern about the racial categories being used. Discussants preferred an inclusive approach, i.e., one in which individuals reporting themselves as both Black and American Indian would appear in both categories and in a combined Black/American Indian category. There was some unease about the use of an "other race" category. Members suggested that data users want tables and charts that "add up" to 100 percent, that persons reporting two or more race categories should not be reclassified into a single race category, and that the Bureau should not report very small race combinations.

In response to members comments, Dr. del Pinal said that the concept of inclusive distributions was interesting, but would create large numbers of tables, and Dr. Tucker noted that the OMB and Census Bureau are examining racial and ethnic reporting in administrative records and evaluating possible changes to the report forms.

Develop recommendations and special interest activities. The four subgroups met separately for additional discussions and to develop recommendations. (See App. A for the specific recommendations and the Census Bureau's responses.)

Closing session. Spokespersons for the professional associations reviewed the recommendations under consideration by their respective subgroups. Members pointed out that the American Statistical Association and the Population Association of America subgroups had made divergent recommendations on tabulating race and ethnic data and suggested that representatives of the subgroups work out any conflicts before finalizing the recommendations. With regard to the agenda for the next meeting, the Committee as a whole suggested sessions on (1) the Bureau's work on the information sector of the economy and database linkage programs, (2) Census 2000 activities, (3) sampling, (4) issues and software for the Post Enumeration Survey and Integrated Coverage Measurement programs, (5) the household file for Census 2000, and (6) small-area estimation for poverty.

There were no public comments.

The meeting adjourned at 12:25 p.m.

Introductory Remarks

Following an introduction and description of the experience leading him to the Acting Director's position, Mr. Holmes made the following announcements:

- Robert Shapiro, the Under Secretary of Economic Affairs at the Department of Commerce had completed his confirmation process and had recently been sworn into office. Mr. Shapiro would be present at lunch to talk with Committee members.

- Ms. Catherine Miller recently joined the Decennial Management Division as head of the communications staff. There will be a number of questions and requests for information, specifically from the House of Representatives oversight subcommittee and the Monitoring Board, for which Ms. Miller will ensure the continuity of the information being provided.

- Fernando Armstrong was appointed the Regional Director of the Philadelphia Regional Office in March 1998.

- The Census Advisory Committee of Professional Associations, the Race and Ethnic Advisory Committees, and the 2000 Census Advisory Committee will hold a meeting on June 3, 1998, concerning racial tabulations for Census 2000. The Census Bureau has brought these Committees together to provide advice on the tabulation and presentation of data on race following Census 2000. The tabulation and presentation of these data must stay within the framework of the Office of Management and Budget's October 1997 decision concerning the standards for maintaining, collecting, and presenting Federal data on race and ethnicity.

- State Action Committee meetings will be held in each state, beginning in the Fall of 1998, in Washington DC. These meetings will bring together local, state, and tribal government officials to discuss Census 2000 outreach and promotion efforts.

- A workshop was held on March 25, 1998, concerning the American Community Survey. People involved with analyzing this survey's data were brought together to discuss research issues and provide preliminary research results. The participants sought ways to improve the American Community Survey and answer questions from data users, including representatives from the Congress.

- The Executive Information System (EIS) is being developed to provide an executive information summary of the Census 2000 Dress Rehearsal. When completed, this system will be available by accessing the Census Bureau's Internet site.

- On March 25, 1998, the Census Bureau's Fiscal Year (FY) 1999 budget appropriation hearings were held. The following day, the House of Representatives Committee on Government Reforms and Oversight held its inaugural hearing on the Census 2000 Dress Rehearsal.

- Two lawsuits have been filed challenging the use of statistical sampling in Census 2000. The first suit was filed in the United States District Court in the Eastern District of Virginia. Representatives filed a second suit in the United States District Court in the

District of Columbia. Both lawsuits seek to bar sampling and the use of statistical methods for apportionment in the House of Representatives. A number of parties have filed motions to intervene on the Census Bureau's behalf.

- The General Accounting Office has completed the first round of an assessment on Census 2000 Dress Rehearsal activities. This assessment included an examination of operations, development of the address list, outreach and promotion efforts, recruiting and retention, sampling for nonresponse follow-up, and the Integrated Coverage Measurement (ICM). The Office of the Inspector General (IG) is also conducting an assessment of Census 2000 operations.

- On October 26, 1997, President Clinton signed the bill authorizing the FY 1998 appropriations for the Census Bureau. These appropriations included $137 million for salary and expenses and $556 million for periodic censuses and programs. From the appropriations for periodic censuses and programs, $398 million has been earmarked for decennial census programs.

- The next meeting of this Committee is scheduled for October 22-23, 1998, at Census Bureau headquarters.

Dr. O'Hare of the American Statistical Association (ASA) subgroup asked to whom the pamphlet, "The Census Bureau: More than Just Numbers" was being sent. Mr. Holmes said this pamphlet was being distributed to census stakeholders. The Bureau does not have a mailing list for this pamphlet.

Dr. Voss of the Population Association of America (PAA) subgroup asked when the State Action Committee meetings were going to begin and if these meetings shared a common goal. Mr. Holmes said there will be a variety of activities planned for these meetings that are associated with the partnership program, including the State Action Committee meetings. The Census Bureau believes it is important to inform the people involved with the census at the state level of the Bureau's progress and what the states can do to make the census successful. The State Action Committee meetings are similar to the Regional Elected Officials' meetings held in 1990. Messrs. Turner and Bounpane are the two contact people at the Bureau for these meetings.

Dr. Scherer of the American Economic Association (AEA) subgroup asked about the legislation allowing the lawsuits filed against the Census Bureau to move on a "fast-track" to the Supreme Court. Dr. Holmes said these lawsuits would move from the District Courts directly to the Supreme Court. There has been a continuing debate about how this process would work; lacking a legal background, he was not prepared to provide a more detailed description of this process.

Census Bureau Responses to Committee Recommendations/Report on October 1997 Meeting

Dr. Scherer of the American Economic Association (AEA) subgroup asked for questions and/or comments on the Census Bureau's responses to the Census Advisory Committee of Professional Associations' recommendations.

Dr. Bell of the American Statistical Association (ASA) subgroup brought the Bureau's attention to an incorrect response to recommendation 16. In addition, he noted that following the October 1997 meeting, the American Statistical Association subgroup and the Population Association of America had submitted conflicting recommendations. After the meeting, the recommendation was reworded so as not to be in conflict. It appears that the Bureau responded to the older version of recommendation 19. Dr. Klerman of the Population Association of America said his subgroup will reintroduce the correct recommendation during the closing session of the advisory committee meeting.

Dr. Klerman said the Bureau's responses to the Population Association of America (PAA) subgroups recommendations, aside from the inclusion of the older version of recommendation 19, were acceptable.

Dr. Scherer said that at the last meeting, the AEA had discussed the evolution of data sharing. At the time, the Office of Management and Budget (OMB) was studying the feasibility of data sharing. In addition, there was legislation concerning data sharing by statistical agencies. He asked the Bureau to update the Committee on this legislation and on any decisions OMB had made.

Ms. Schneider said there were no formal actions taken on the legislation. There were some hearings and testimony offered. The Census Bureau's concerns regarding the legislation had been brought to the attention of the Congress through the Administration.

Dr. Knickerbocker said that new legislation has been introduced since the last meeting— Senate Bill S.1404. Section 1 of this legislation is an expression of the interests of Senator Moynihan. A commission would investigate the advantages and mechanics by which major parts of the statistical community could be consolidated. The commission would focus specifically upon the Census Bureau, the Bureau of Economic Affairs (BEA), and the Bureau of Labor Statistics (BLS). The legislation implies that consolidation is the objective. Section 2 is another variant on data sharing. It is an expression of Congress' sentiment that data sharing is a good thing. Since there is controversy surrounding data sharing, the OMB is waiting for the administration to provide details on how data sharing will be implemented.

Independent of this data-sharing legislation, the Census Bureau, the BEA, and the BLS have been meeting to sort out the mechanics and implications of data sharing should it become a reality. There are complications that must be addressed before data sharing can become a reality. For example, the BLS relies upon data collected by states to operate its employment, wage, and earnings programs. The BLS returns certain data back to the states. Some of the information is considered confidential, including industrial classification data. If data sharing were implemented, the BLS would have total access to the information on the Standard Statistical Establishment List. It is uncertain to what extent these data, protected by title 13, will "leak" back to the states.

Mr. Scherer asked if S.1404 would require an outright physical consolidation of the statistical agencies. Dr. Knickerbocker said that there would be an organizational consolidation. He said he believed there was bipartisan support for S. 1404.

Mr. Adams of the American Marketing Association (AMA) subgroup said he would like to see estimates of the cost to implement data sharing and what the Census Bureau's position is

on that issue. From a marketing view point, data sharing could have a negative effect upon response rates, since it may be seen as an infringement upon the public's privacy and right to confidentiality.

Dr. Knickerbocker said the greatest opportunity to save money would be through a consolidation of the Census Bureau and the BLS, specifically in sharing information on industrial classification. Both the Census Bureau and the BLS collect identical economic data from businesses and both maintain an industrial classification list. Consolidating the two agencies would obviously save money.

Ms. Becker brought an article to the attention of the Committee that concerned data sharing. This article questioned whether Employment Survey 202 data (administered by the BLS) belonged to the states or the BLS. Some states have concerns about data sharing because some state laws do not permit the sharing of certain data.

Dr. Scherer asked how the implementation of the North American Industrial Classification System (NAICS) was progressing. In addition, he asked if there had been any movement toward giving the NAICS information sector more attention. Dr. Knickerbocker replied that Mr. Mesenbourg would address NAICS during the American Economic Association subgroup meetings. The first appearance of NAICS in census products would occur in January 1999. These products will be the initial data offerings from the 1997 Economic Census.

The Census Bureau's Plans for Poverty Measurement

Dr. Weinberg briefly updated the Committee on the progress of revising the Standard Occupational Classification. Approximately 800 detailed occupations have been identified. These are aggregated into 98 minor groups and 23 major groups.

Following a discussion of "The Census Bureau's Plans for Poverty Measurement Research," Dr. Lillard of the American Economic Association (AEA) subgroup said he believed the move from the Current Population Survey (CPS) to the Survey of Income and Program Participation (SIPP), as recommended by the National Academy of Sciences' (NAS) panel, was a good move. The SIPP would make information available that would make the analysis of poverty and other substantive issues better. The data provided for households will provide an opportunity to study data in terms of attrition.

Dr. Lillard said one of his major concerns was panel attrition. It is important to follow individuals even if the households or family units split-up. While this will help reduce attrition, there may still be some attrition among individual family members. He suggested that initial refusals in addition to panel drop-outs be studied. If initial refusal is an indication of being in poverty, he suggested a follow-up be conducted to contact refusals.

Dr. Weinberg said the Bureau has undertaken a special study of the 1996 SIPP panel's first wave nonrespondents. Mail response to the follow-up survey was about 20 percent. During the SIPP's first wave, the Bureau also studied the impact of incentives on response and found that nonresponse was reduced among low income households when incentives were offered. Additionally, the Bureau is testing a booster-shot incentive in wave 1 to determine what impact this will have on attrition.

Dr. Lillard said that if low-income families exit at higher rates early in the sample, and if full year's income is being studied, then attrition becomes a significant issue over the 3-year study period. This problem may be resolved if the constant attrition biased design were used. The constant attrition bias design measures year-to-year changes accurately; however, changes in poverty will change attrition.

Dr. Weinberg said that attrition bedevils all longitudinal surveys. He understood that changes in poverty may affect attrition. The Bureau has been working to decrease nonresponse for all surveys. He added that some national population surveys are reweighted, but this does not correct for unobserved characteristics.

Dr. Lillard said exclusion of medical expenditures and medical insurance premiums from income measurement and the threshold was a issue going to the heart of the measurement and meaning of poverty. He was uncertain how the separate medical risk index would be combined with or be made separate from the poverty issue. Medical care can be as essential as food and housing, but it is much less stable. It is different from housing and food in that medical expenditures are more volatile over time. If a family member has a chronic condition, out of pocket expenses can be skewed, affecting the availability of other resources.

Dr. Lillard suggested that the Census Bureau consult a paper on the burden of prescription drugs and the elderly and other works showing that the margin of consumption is higher when an individual is in poor health. In addition, he suggested an experimental supplement be conducted that would implement a "Consumer Expenditure Survey (CES)-type" survey in addition to the SIPP. If the two were conducted together, or at least had a large enough sample or subsample, it may be possible to address some issues more directly.

Dr. Weinberg said it would be difficult to supplement the SIPP with the CES, since respondents have indicated that these two surveys are the most burdensome. The Bureau will use statistical matching to supplement the SIPP.

Dr. Lillard said the implication of cohabiting couples and housemates who share expenses is important, since household composition is a factor when conducting poverty analysis. His own work indicates substantial turnover among cohabiting couples. Within a year, households within the study sample experience divorces, births, marriages, and movement of children in and out of the household. These changes are important considerations when trying to measure yearly income.

Dr. Ghosh of the American Statistical Association (ASA) subgroup said he was concerned about biased estimates. Some kind of model based analysis has to be done to account for attrition bias to produce yearly estimates. He was pleased that three panels would operate simultaneously. Other than the bias reduction, the measure of accuracy for these estimates is an important issue. He asked if there was a better method than the CPS. Dr. Weinberg said the effectiveness of another method depended upon what was being measured. Dr. Ghosh said that even after moving from the CPS to the SIPP, the CPS could still be used for comparison and as an auxiliary measure for producing estimates.

In response to Dr. Ghosh's request for clarification, Dr. Weinberg said medical risk would be a separate measure used to tabulate a category. Discretionary medical expenditures, like cosmetic surgery, would be difficult to separate from medical out-of-pocket expenses.

Dr. Ghosh said he had some concerns regarding the Consumer Expenditure Survey. He believed that a 3-year sample would reduce, but not eliminate bias in that sample.

In response to a question by Dr. Ghosh, Dr. Weinberg said there would be a lower poverty rate, e.g., if there were a lower threshold in Mississippi. Dr. Ghosh urged that geographic variation be considered, otherwise a comparison of poverty rates from state to state would not be fair.

Dr. Ghosh said treating cohabiting couples like married couples may not be problem, since they experience similar economies of scale. He noted that new research on the extent to which housemates and roommates share living expenses was a good topic of research; however, he would doubt the accuracy of the data. In his experience this information is often untrustworthy.

Dr. Weinberg said the March 1998 CPS asked households to report their income in 1997. However, the composition of these households may have been different in 1997, so some inaccuracies are introduced. The SIPP poverty data shows family composition and income from any 1 month, and the Bureau aggregates the income for 12 different months. If family circumstances change, the appropriate thresholds are taken for each month and also added up over the year. One advantage of the SIPP is that its data are closer to the circumstances than the subsequent March.

In response to a question by Dr. Ghosh, Dr. Weinberg said that there may have been some confusion over the equivalence scale. Orshansky's "three-times food measure" will not be used. Instead, a formula proposed by the National Academy of Sciences' panel is being evaluated. Shifting the equivalence scale will provide different distributions of poverty.

Dr. Myers of the Population Association of America (PAA) subgroup said that after reading Dr. Weinberg's paper and learning of the burden placed on the Census Bureau, he realized the data problems this "can of worms" represents. We can no longer use the decennial census for poverty measurement because there are too many variables lacking. He believed there were three issues associated with poverty—(1) the level of poverty, (2) the trend, and (3) the inter-area differences.

Level of poverty, by itself, is a political barometer. It can only be judged on intuitive, reasonable grounds. As Dr. Weinberg's paper emphasized, the $3,000 target number, used to judge whether Orshansky's measure was reasonable, was an arbitrary number that people seemed to agree on. That kind of benchmark will be essential "underneath" all this detailed statistical analysis. If the debate over the revision of the poverty thresholds were not occurring at the same time as the Census 2000 debates, poverty threshold would be getting more attention. With the geographic differences being important for the allocation of revenue based on the poverty variable, defining poverty has consequences similar to the measurement issues associated with Census 2000.

Dr. Myers believed switching from the CPS to the SIPP was a big deal. The SIPP's measurement of the poverty level is one-third lower than the CPS's. The numbers may have to be adjusted, which may raise questions. The SIPP's poverty measurement may be lower due to the differential attrition, or it may be a result of definitional issues. Switching from CPS to

SIPP not only results in a difference in level, but also adds problems associated with panel structure and attrition.

Dr. Myers said he would be interested in learning how the Bureau plans to interface experimental measurements with the other data series, including the decennial census, the American Community Survey, and the Small Area Income and Poverty Estimates program. Ultimately, these will all have to be cross-referenced. The national level data series are needed to get the levels right and possibly the trends. Ultimately, the geographic differences will need to be studied. This will be a problem, since the poverty rate has been most erroneous over the past years.

Dr. Myers added that he often compares Los Angeles to Mississippi when discussing poverty, since Los Angeles does not have rent control. The inequities involved with using the same poverty levels throughout the country are severe. He was confident that Dr. Weinberg's housing experience would enable him to adjust for geographic cost-of-living.

In response to Dr. Myers' question, Dr. Weinberg said the time horizon for the new SIPP data on poverty would be sometime in the year 2003. He hypothesized that it would likely take about 10 years before changes that took into account all the evaluations were incorporated. Currently the statistics from the March CPS are published during September. The last SIPP data-collection month for 1997 was April. The SIPP's longitudinal processing, which fills in missing data, and performs edits for consistency, is a more complicated process than the CPS's. An additional 4 months will be required even before processing of the 1997 data are begun to allow longitudinal editing. If model-based estimates are added rather than direct-based estimates, additional time will be added.

Dr. Short said she has looked at some of the imputed values on the CPS and compared them to the values on the National Medical Expenditures Survey. She said she is surprised how close the distributions of medical expenditures are; however, without a blackboard, she would have difficulty giving any further details.

Dr. Stolzenberg of the PAA subgroup said he agreed that the subjective nature of the variables had to be recognized. Past research has shown that people often define "poor" as one-half the median income. The problem facing the Bureau is not making objective measurements, but keeping the debate about these measures from damaging the Bureau's core activities. Additionally, poor people tend to be impoverished in some ways and not in others. For instance, obesity is most prevalent among the poor. They are the most "caloric" overfed population in the country. A study of calories is important when discussing poverty.

Dr. Wilson of the PAA subgroup said that the Committee did not adequately address attrition bias and how the use of the SIPP panel will raise concerns. He was pleased that Dr. Weinberg's paper addressed this issue. Regarding the CES, the National Academy of Sciences' panel recommended using the survey as a basis for estimating the poverty threshold. Dr. Weinberg's paper expressed some concerns about the quality of these data when this survey was used. The panel suggested the Bureau of Labor Statistics try to determine if it would be possible to make some changes in the CES—making it a more effective tool for obtaining the basic information needed to estimate the poverty threshold. Dr. Weinberg said these is a plan being developed to expand the CES sample by 50 percent in urban areas. This

expansion will improve future poverty threshold data. He added that the increased sample in urban areas would still not provide enough data to create national poverty statistics.

Dr. Scherer of the AEA subgroup said mobility also is an important aspect of poverty. Medical expenditures vary so widely that the suggestions to break out a separate indicator of how medical expenditures are affecting the status of the poor is a very good one. In regards to Dr. Stolzenberg's comments on the caloric intake of the poor, it is not until one studies the very rich that an improvement in the quality of food consumed is observed. Additionally, clothing does not necessarily vary with income; however, housing is a kind of residual. As rents increase, the housing quality of the poor decreases. It is important to include the quality and cost of housing among the poor.

Dr. Scherer asked who would be making the decisions regarding the poverty threshold and income measurement. Dr. Weinberg said that the Office of Management and Budget (OMB) has emphasized the importance of the agencies making technical, not political contributions. There is a separate group established in the Executive Office of the President to offer policy guidance. The final decision will be made by the OMB. It will be years before changes could be made. Anyone who could be affected by the decisions will probably be afforded an opportunity to comment.

Ms. Becker of the Population Association of America (PAA) subgroup said that housing quality can not be measured. The market of any given geographic area (whether there are housing shortages or surpluses) makes a big difference. She believed a lot of the relationships at the bottom end of the economic scale are going to change. As people leave welfare and move into the workforce, they are likely to disperse geographically. The economy also will have an impact. Welfare reform is going well now because the economy is good. If there is a recession, reform may not be as successful. She urged the Bureau to make a special effort, within the context of existing studies, to separate and measure what is happening demographically and socioeconomically as people within sample move off welfare. It should be recognized that the total number of people in poverty and the level establishing poverty will probably be established politically or policy-wise, not technically. She added that she had attended a meeting recently where one of the speakers said that Orshanky's $3,000 level was adopted because it fit peoples' sense of what being poor meant.

Dr. Willis of the AEA subgroup said the two largest and most obvious policy issues are those involving children. The next issue is the aging of the population, measurement of poverty amongst the aged, and the role of medical care and expenditures. It is known that there is a high correlation between mortality rates, health, and education. Including health in the poverty measurement will be important. The scale of the economy should also be taken into account as well as the extent of single-person households and two-person households, including a caregiver who provides time and/or income. Assets can also be used to provide income, especially among the elderly.

Dr. O'Hare of the ASA subgroup asked if there had been any progress on the proposal to add questions to the long form. Dr. Weinberg said that the questions were presented to the Congress on March 30, 1998. We proposed no additional questions would be added to the questionnaire.

Dr. Maynard said that as an analyst, she was unsure how to allocate an effort to worrying about precision in measurement in the area that is relevant to measuring poverty versus getting an accurate income and need measure for everyone in the income distribution. It seemed that a system had been set up that was worrying about accuracy of the measures of every individual when we could concentrate a lot of the precision work and estimation down in the lower end of the at risk population.

Dr. Weinberg said the design of the SIPP had been changed so there would be an over-sample of the low income population, since this group is the primary focus of the survey. There is a pretty large constituency that does not want the SIPP to simply be a measure of just the low income population.

Are We on the Right Track with the Corporate Marketing Program? (AMA)

Since most of this 2-day meeting was devoted to the Census 2000 Dress Rehearsal and its measurements, Mr. Adams said he and the other members of the American Marketing Association (AMA) subgroup would like to know about the Bureau's promotion and marketing programs for the dress rehearsal as they are being implemented at the test sites. The subgroup members were especially interested in the confidentiality issue and how the agency was approaching it. Mr. Wynegar said that could be arranged. Mr. Kavaliunas suggested showing a video tape that covered commercials developed by Young & Rubicam.

Mr. Wynegar expressed appreciation for the subgroup's ongoing guidance, advice, and continuing support.

Dr. Spiro, referring to the Bureau's response to the subgroup's comments on data sharing, said there was a sentence suggesting that "there is a paper being prepared..." that did not reflect very good marketing strategy. She felt that the way it was said would leave a bad impression, especially on legislators, when in reality there was no paper being prepared. She also noted that the end of that response included a statement "...in the end, the choice to share data is ours" after going on and on about how the agency would look into the costs and benefits. She did not think it was a very politically astute way to say that. She was afraid that people would read the language of the response as if the Bureau did not care what others suggest—it would do whatever it wants.

In reference to the AMA subgroup's recommendation no. 15 from the previous meeting on FEDSTAT one-stop shopping, Dr. Spiro said the subgroup recommended direct mail, placing advertisements in certain publications, etc. However, the Bureau's response was that there was not enough money to do so. She wondered why the agency did not know that there was not enough money, and why it was not aware of its budget constraints. It would be nice if the subgroup knew ahead of time that there was not enough money for an advertisement campaign; then the subgroup could have formulated its suggestions differently and more creatively.

Mr. Wynegar thanked Dr. Spiro for her overall point of view and apologized for not communicating the constraints and the assumptions the Bureau was using in its planning for the Corporate Marketing Program to the subgroup adequately.

Mr. Adams admired the Bureau's corporate identity and the logo. He asked Dr. Etzel to comment on the Bureau's Corporate Marketing Program.

Dr. Etzel submitted his comments in writing (see handout). His first observation was that new products were relatively attractive when they first become available, but their attractiveness declined over time. For these products, a marketing program needs to be developed that would make them more attractive later in their lives and maintain a revenue flow.

Mr. Wynegar agreed that that was the secondary, if not primary, purpose of the Bureau's program. With its Corporate Marketing Program, the agency is trying to provide a test of a process, i.e., targeting specific markets in an organized way using a small-scale pilot project.

Dr. Etzel commended the notion of targeting and identifying market segments, but given the size of the budget and the markets, they were not segmented enough. The ability to segment these markets more specifically depends upon the availability of data. He asked if the Bureau had enough information to do so.

Mr. Kavaliunas asked Mr. Thompson if he had looked into the additional segmentation of financial institutions. Mr. Thompson said he did and explained how it was done.

Dr. Etzel noted that there were two separate issues—(1) who to target to send the communications materials (e.g., sending a piece of direct mail to a corporate director or to a research director, etc.) and (2) the segmentation issue. Within the context of the 488 commercial banks that bought census products in the past, the Bureau needs to find out what attributes of a bank might be the reason for it to buy census data in order to segment the banks. The size of a bank is not important for segmentation.

Given the Bureau's budget constraints, Dr. Etzel did not think that it would make any sense to do promotion to 8,586 commercial banks; it would not work. However, if the agency targeted about 500 banks out of the 8,586 based on their attributes, the test would provide meaningful data.

Focusing on the promotional components of the Bureau's plan (i.e., getting the message out that the product is available), Dr. Etzel believed the promotional material for old data should emphasize that the price has been lowered. Mr. Kavaliunas added that some of these old products, especially the geographic products, have been value-added and updated and are being promoted as "new and improved."

Dr. Etzel added that the rationale for the expected return for investment should also be considered before a direct mail survey. Mr. Wynegar said the marketing survey averaged about 2- to 3-percent response rate.

Dr. Etzel did not feel the Bureau's marketing survey's primary message ("being the official source of government statistics...") was exciting, nor was it a "benefit" message. He also added that appropriate benefits messages can be developed only after targeting and segmenting the market.

Mr. Kavaliunas said the banks were a captive audience because they are required to report mortgages by census tract, and the Bureau provides that official statistics. Dr. Etzel asked why only 3– to 4–percent of the banks buy the Bureau's product. Mr. Selby said that was because banks were buying the data from other intermediaries. In that case, according to Dr. Etzel, the Bureau needs to show why they should buy the product from the Bureau.

Ms. Stershic said the Bureau should find out if the banks were getting assistance from the intermediaries in how to use the data. Dr. Etzel agreed that the intermediaries make their data more user-friendly by providing assistance in how to use it.

Overall, Dr. Etzel did not believe that a single exposure (i.e., a one-time survey or program) would be very effective. He expected different outcomes from different media; he also believed that if the segmentation issue could be refined, everything else would fall in place. Considering the budget constraints, he suggested using a very small subset of a group instead of all four groups (mentioned by the Bureau's Marketing Services Office staff) to get some meaningful results.

Dr. Spiro agreed with Dr. Etzel on targeting and segmenting the universe of banks. In addition, she noted that the Bureau's plan mentioned about the size of the banks, but not about the attributes of the banks—e.g., some banks do more commercial loan business, some are more into other specific areas, etc. Attributes would be more useful in segmentation than the size of banks.

Mr. Wynegar apologized for not communicating to the subgroup what the agency had done in terms of segmenting. He noted that the Bureau had done quite a bit of research on further segmentation of the financial institutions.

Dr. Spiro said she knew the Bureau had done focus groups and surveys on why people buy certain products, but that was different than segmenting on demographic and other characteristics of organizations. Mr. Wynegar said the Bureau had done that to some degree.

Dr. Etzel asked what type of data the Bureau had on the banks. He also felt that there was a need for the Bureau to find out more about the banks—i.e., at least two or three common attributes of the banks who purchase census products—and then target them.

Ms. Dickinson said the Bureau classified the banks by sites versus numbers and observed whether or not the branches had the ability to purchase data independently. The agency also had a conference call with the First Union Bank to see how its operations worked. Dr. Etzel said it was an excellent idea, but the Bureau should not want to generalize from one bank. Dr. Spiro suggested the Bureau also may want to find out which banks have research departments.

Mr. Selby summarized the key findings of the Bureau's Corporate Marketing Plan focus groups and distributed a handout listing the 11 findings and some product promotional and packaging materials used by the Bureau. He noted that the Bureau did not address the benefits message as being the main goal during these focus group sessions. In summary, the agency was perceived as having integrity and doing a good job in collecting information, but not in packaging and distribution of the information. Most of the current data users were unaware of the number of products and services that the agency provides. For the Summary Tape

File 3, some of the users were willing to pay three times more than the actual cost. The Bureau should promote itself more to the public as an authority in demographic information and use a variety of marketing venues. Data users did not want 1990 census data in 1998. The Current Population Survey data is poorly organized, its tables are not easy to follow or aesthetically pleasing, and the information was past due and not accurate. The Bureau should give the users the ability to tabulate data online and needs to develop user-friendly products. Agency-wide customer service should become a priority, especially technical support for products. The Bureau needs to have a brand new campaign, like the one carried out by the U. S. Postal Service, to establish itself.

Mr. Wynegar said the Bureau did an online survey on its Internet site; 21-percent of those who completed the survey bought census products because of their visibility on the Internet site. Dr. Etzel said that the information was very valuable to the Bureau, but he had some reservation in putting more and more money into it. Mr. Adams noted that online is not just a marketing tool, it is the lowest-cost way of providing service.

Ms. Stershic said the Bureau ought to start relationship marketing strategies, which she did not see mentioned in the agency's plan. For segmentation, the Bureau should start to work strategically on building relationships with the users to be able to sell more products. She suggested having customer roundtables that would not cost anything, newsletters, direct mail, seminars, and so.

Ms. Ashcraft concurred with other members' comments on looking at the subgroups, segmentation, and building relationships, and agreed that direct mailing was a good opportunity to do so. She noted that, in previous AMA conferences, the Bureau's exhibit booths were staffed by only one person. There were a lot of publications to look at, but no excitement like some other booths had in the form of contests to try to get people to come to their booths. She believed the agency now can show people how to go online for census data or how to use census CD-ROMs to make the booths more interactive and exciting.

Dr. Spiro gave some information to Ms. Ruffin on the American Bankers Association, the Association of Credit Unions, and the Association of Investment Management and Research corresponding to the Bureau's different target markets. She also gave the name of a professor at the Indiana University's School of Business—Dr. William Sartoris—who might know which of the above associations the Bureau should target. She noted that the Bureau's use of the terms "useful" and "easy to use" as confusing, because they convey two different meanings.

On segmentation, Mr. Adams agreed with Dr. Etzel and suggested at least one follow-up mailing after a direct mailing campaign even if it is to a subsegment of the total sample. In terms of the sales messages, he noted that, as mentioned by other members of the subgroup, the term "easy to use" has a lot more leverage than the terms "official statistics" or "useful." He showed a chart describing the elements of the persuasive message and related the information to what the Bureau has been trying to do. He noted that the U. S. Census Bureau logo was sufficient for data users to recognize that the product came from there and, therefore, there was no need for the message "official statistics." Mr. Adams also emphasized that if old products are value-added, the Bureau must clearly label them in the promotional materials as "new and improved" or so. At the next AMA conference, he suggested the Bureau demonstrate online activities with census data.

In reference to segmentation, Dr. Etzel gave an example of a catalog retailer specializing in large-size clothing. The company segmented the market for people who buy large-size clothing and created a database of 21,000 customers, then subsegmented that into 75 segments according to demographic characteristics and purchasing behavior.

Dr. Spiro asked if the Bureau sent forms to banks for reporting mortgage information, and, if so, a survey easily could be attached to the form. Mr. Kavaliunas said no; the banks used census maps and products to report mortgage information to the Federal Reserve Board which has its own online facilities.

Mr. Selby noted the development of the Bureau's Landview III software for the TIGER (Topologically Integrated Geographic Encoding and Referencing) system in conjunction with the Environmental Protection Agency (EPA). This software shows on maps EPA "brownsites" that developers are legally responsible for cleaning up if they purchase that property, and the banks could be interested in having this information. He went to the American Banking Association convention to talk about this software, but no one was interested in knowing about the Environmental Protection Agency "brown sites."

Dr. Etzel said, while considering a new product, an organization needs to talk to the market to find out if this would be a viable product before developing it.

Dr. Spiro said the Bureau should not feel bad, because 50 percent of the new products developed each year fail.

Mr. Kavaliunas noted that the Bureau has been preparing product profiles for the new products that are being released. Some of these profiles are distributed as handouts and some are used in direct mailing depending on the product and the budget.

Mr. Selby said the Bureau mailed out 10,000 CD-ROM catalogs in March 1998 to foreign college, university, business, and research libraries, based on the Library of Congress' experience of having $7 million worth of business each year with the libraries. He distributed a schedule (listing meeting dates, data products, etc.) that was given out to users at multiple international trade conferences attended by the Bureau's Foreign Trade Division. He noted that this schedule got a great reception from people.

Dr. Spiro asked if the information on the Bureau's CD-ROMs was the same as the information available on the foreign trade Internet site. Mr. Kavaliunas said some of it was the same. Mr. Wynegar added that the agency has a subscription service on the Internet, part of which is a foreign trade data base.

Economic Census Update (AEA)

Mr. Govoni said much of his presentation will be on the reasons for total and partial nonresponse to the 1997 Economic Census, and what is being done to affect this nonresponse. The Bureau carried out the initial mailout of approximately 5 million economic census questionnaires in December 1997, and by the end of February 1998, processing was underway at the Data Preparation Division's office in Jeffersonville, IN. The Bureau plans to complete data collection by the end of September.

Mr. Govoni noted that the unit response rate to the census so far has been disappointing—running about 3 percent short of the 1992 pace. He suggested that the final unit response rate would be about 80-85 percent, which would represent a 90-95 percent coverage in terms of total receipts. The difference in the item response and coverage is the result of the fact that response is skewed because large companies both respond better than small companies and are followed up more intensely than small operations, and because the variable selected for comparison—receipts—is one that is very well reported. The Census Bureau has encountered differential response to its surveys and censuses by item, size of company involved, and other characteristics. This differential is the result of the Bureau's collection strategy (e.g., intensive follow up for large companies), reporting patterns (response rates differ by size of company and even by geographic region of the country), and data cleaning and imputation strategies.

The Bureau's collection strategy focused on 5 million economic establishments, which received one of 475 specialized census questionnaires. The Bureau used over 200 long-form questionnaires for collecting detailed data for the manufacturing sector alone, and employed short-form questionnaires asking for limited and summary data to reduce response burden for small companies. Since the detailed data are not requested on the short forms, the Bureau estimates totals for those detailed items using industry averages. To accommodate the adoption of the new North American Industry Classification System (NAICS), the Bureau also mailed approximately 1.5 million classification forms to economic establishments that normally would not receive an economic questionnaire, while basic summary data for about 14 million nonemployer businesses were taken from administrative records provided by the Internal Revenue Service (IRS) and Social Security Administration.

When all of these sources have been combined into the economic census data file a casual user might think that everything has been covered and reported in the census.

Regarding reporting patterns, Mr. Govoni pointed out that response is not the same for all items requested. Respondents do not maintain information in their own databases in the detail and form most convenient for economic census data collection and processing. Employment and payroll, and to some extent, receipts, are generally fairly readily available. Other subjects are not as easily covered; e.g., most respondents have had considerable difficulty reporting the cost of materials consumed, with many lacking the detail the Bureau requests. These data are critical to the Bureau of Economic Analysis's (BEA's) input/output table calculations, so the census must try to collect as much detail as possible. Response by type of activity also varies, with apparel companies being particularly poor reporters. In geographic terms, Midwest companies are among the better reporters.

Mr. Govoni said the objective of the Bureau's data cleaning activity is to ensure that the census has valid data. After the data are keyed, the Bureau performs a series of computerized checks to identify major problems with the data—e.g., the agency sums up the establishment data reported by a company, and compares them to the the data supplied to the IRS. If there are discrepancies the most likely reason is that there are establishments that are not being reported, or are being overreported, and the Bureau contacts the company to confirm the data supplied. The processing operation also validates the individual respondent's industrial code and scans each record for missing critical items.

The Bureau has developed a new complex edit and imputation system—referred to as "Plain Vanilla"—to provide a simplified edit and data imputation system. The system relies on reported payroll and a series of interrelated ratios to use to impute missing data items for each record (i.e., the system assumes that a company in a given industry, with a specified payroll, will share certain other characteristics with other companies of the same type and payroll size). The Bureau expects to refer 20– to 25–percent of economic census records to analysts for review.

Once the individual records are edited, corrected as needed, and data imputed where needed, the data from the records are tabulated and undergo aggregate review before the first statistical reports are prepared and issued. The Bureau plans to issue an advance national level report—showing data under both the NAICS and the Standard Industrial Classification (SIC) classification systems— early in 1999 as the first product from the 1997 census; this will be followed by approximately 500 industry and state reports (these reports will classify industries and activities on a NAICS basis only, with bridge tables providing continuity with earlier census data published using the SIC code). Some additional subject series reports will be issued in 2000 and after. The Census Bureau gives the highest priority in its aggregate review to those data needed by the BEA for its activities and to trade associations, then to census-to-census comparisons, and then to large records that will have the most impact on individual cell totals. He noted that the Bureau has very sophisticated tools for aggregate review, but overall resources to support the activity are limited, which means the review has to focus primarily on large cases with the most impact on the data. As a practical matter, approximately 70 percent of the review time for the aggregate data is spent on product review.

In response to a question by Dr. Scherer, Mr. Govoni said company downsizing over the past few years has included cuts in staff assigned to respond to government inquiries, but he believed this situation was now stabilizing.

Responding to a comment by Dr. Knickerbocker, Mr. Govoni said he would like to see the 1997 Economic Census advance report released by January 1999, but was not certain that goal could be met.

Replying to a question by Dr. Scherer, Mr. Govoni said the Bureau usually assigns economic census analysts by 2-digit classification groups—one analyst per 2-digit group—so each analyst may work with dozens of 4-digit industries.

Dr. Dunne said he put together some analysis of response to the 1992 Census, concentrating on shipment data from the 1992 Census of Manufactures. He looked at how response varied by industry, size of company, whether the company was a single- or multiestablishment operation, and whether the company was a new establishment or was nonrespondent in a prior census. He noted that, compared to the largest establishments (i.e., those with 250 or more employees), the smallest establishments (1-10 employees) had a 20-percent lower response rate. New establishments are about 12 percent less likely to respond than older ones, and, because they tend also to be smaller operations, overall response from small new companies approached a low of as little as 60 percent.

Dr. Willis pointed out that in household surveys, researchers have found that asking for "global" figures on most financial questions results in under-reporting. People seem to have an easier time supplying the information when asked for details and then sum up their individual

responses. Mr. Mesenbourg commented that in 1987 the Census Bureau conducted a survey to compare responses obtained using the long- and short-form questionnaires, but he did not recall the specific results. Mr. Govoni said the Bureau did a study some years ago on collecting data using short and long forms, and while the totals reported by respondents to the survey were not very different, the response rates obtained by the two form types differed considerably. Dr. Knickerbocker pointed out that the basic assumption made by the Bureau is that smaller companies will find it easier to supply simpler numbers. That may not necessarily be the case.

Replying to a question by Dr. Scherer, Dr. Dunne said that new and small companies had the worst response rates. Single establishment companies are 6 percent less likely to respond than are multiestablishment companies, and single establishments that were nonrespondent in the previous census are 16 percent less likely to respond to the current census. Response declines for those items, such as data on inventories, that are not among the basic questions asked for all companies.

He stressed that the Bureau's basic strategy is aimed at obtaining aggregate data, rather than microdata. Any microdata user must be careful when using economic census data to recognize when they are using reported data and when they may be using imputed data. Mr. Govoni emphasized the importance of users realizing that the Bureau's whole economic census process is geared toward producing aggregate statistics; the linkages established are all designed to aid in the production of aggregates.

Mr. Mesenbourg pointed out that the Bureau does obtain some good auxiliary control variables from the IRS administrative records for small establishments, but these cannot be used for multiestablishment operations.

In response to a question by Dr. Gort, Mr. Govoni said the Bureau tries to identify imputed data in the records, not only for users, but for the analysts' work as well. In the census publications, the Bureau includes imputation rates for industry, state, and even by size of company in some of the tables. The agency can identify the imputation done for an individual establishment.

Dr. Knickerbocker commented that the basic purpose of the economic census is to supply benchmark data for the BEA's calculations. Absent this information from the census, the BEA would be completely adrift in trying to measure quarterly changes in the economy.

Replying to a question by Dr. Maynard, Mr. Govoni said the Bureau uses industry averages from the Annual Survey of Manufactures, previous censuses, or reported averages from specific industries in the census year for its imputation work. He added that the data have to be "clean" to have good imputation factors.

Mr. Mesenbourg commented that the Bureau assumes that respondents and nonrespondents to the census have generally similar economic characteristics, so the question for the agency is, what factors inhibit response? One variable that was identified was the knowledge of the mandatory nature of the census, and the Census Bureau has emphasized that in the 1992 and the 1997 censuses. The effectiveness of this procedure may be waning as respondents realize the Bureau is unlikely to prosecute companies for nonresponse, so the agency may have to start looking for some other way to help encourage response.

Dr. Dunne noted that, for manufacturers, new establishments' investment numbers are proportionately higher in terms of value of shipments than are those for continuing establishments. Using those numbers for new establishments would underestimate the investment that is going on.

In reply to a question by Dr. Maynard, Dr. Dunne said that when the Bureau has payroll numbers for an industry, it can look at the industry for ratios of receipts to payroll. Mr. Govoni added that the actual imputation operation is not as simple as it has been described, and the precise details of imputation will depend on the particular item and industry involved.

How Can the Census Bureau Get Consistent and Useful Feedback From Its Customers? (AMA)

Ms. Dickinson asked the American Marketing Association (AMA) subgroup members for their input regarding her paper, "How Can the Census Bureau Get Consistent and Useful Feedback From Its Customers?" The paper outlines a Web-based proposal for a new, customer-friendly page, the Customer Information Exchange Page, which would include an online newsletter and a regularly scheduled, "live" topical chat room. Mr. Wynegar added that another means of obtaining customer feedback would be for the Bureau to sponsor census user conferences, which it was considering at the time of the previous Committee meeting. Due to staff time constraints, however, the agency is not likely to sponsor any user conferences until after Census 2000. In addition to the Customer Information Exchange Page, he asked subgroup members for comments on the utility of user conferences or any other mechanisms for obtaining user feedback.

Ms. Stershic, Dr. Etzel, and Ms. Ashcraft respectively made the following points in favor of user conferences—

- There is no substitute for the face-to-face contact and relationship-building opportunities provided by user conferences.

- User conferences allow for better networking, not only between the Bureau and its users but among the users themselves.

- These conferences are a better vehicle for "how to" sessions in which case histories can be presented in more detail.

Mr. Wynegar said that the cost of the conferences was not an issue for the Bureau; the issues raised by its executive staff were staff time and value to the agency. Drs. Etzel and Spiro suggested the agency also consider partnering with other organizations, such as the Association of Public Data Users or the AMA, to reach data users at their meetings. Dr. Spiro added that user tutorials could be given at these meetings. Both members cautioned against overexposure. Census users can attend only a limited number of conferences. Having participated in census tutorials at other meetings, they might be unlikely to attend conferences sponsored by the Bureau.

Ms. Ashcraft commended Ms. Dickinson for her paper and involvement with the Customer Information Exchange Page, and addressed the three questions the paper asks—

1. "Is our heavily concentrated effort toward the electronic media for communications exchange appropriate?" Yes, this concentration is appropriate. The Internet is a vehicle for customer communication already and will be used for this purpose more and more in the future.

2. "Are there any legal, technical, logistical, or political issues that we should be concerned with when considering interactive chat rooms?" This is a good question for the Bureau to be asking now. The agency may want to forego "live" chat rooms or opt for a delayed form of response over which it would have more control. Perhaps some form of registration should be required to ensure contributions only come from the desired audience. Closed bulletin boards work well since irrelevant or self-serving responses can be edited out. Mailing lists in conjunction with these bulletin boards can provide those who register (subscribe), with e-mail notification of the other participants' responses. In this way, the subscribers can choose to respond "publicly" on the bulletin board or privately by e-mail. The most important thing for the Bureau to guard against, however, is that the names and addresses of the participants not end up on a junk e-mail list.

3. "What other arenas or forums would members recommend that we investigate that would invite customers to share their ideas and opinions?" The closed bulletin board coupled with a subscriber mailing list.

In response to a question by Mr. Kavaliunas, Dr. Etzel suggested the Bureau consider a listserv, bulletin board system similar to the one used by L-MAR. This system requires a person to monitor up to 1,000 subscribers, and there usually is a 1-day delay before messages and responses are posted. Another advantage of this system is that the monitor can cull messages and responses that are no longer relevant so that the participants do not have to sort through too much material.

Dr. Etzel also suggested the Bureau consider inserting a warranty card with the products it delivers. Mr. Kavaliunas said the agency had sent out a customer-satisfaction postcard 3 or 4 years ago and that it was still receiving about 10 responses a month. Dr. Etzel added that the primary value of these cards was to build customer databases.

Mr. Adams said that online feedback from customers should be kept in perspective and viewed in the larger context of customer-satisfaction surveys. Ms. Ashcraft added that many companies used the same approach to evaluate information from their complaint departments. Ms. Dickinson said the Bureau had obtained feedback from its Web site users a few months ago, which it then profiled against its customer database.

Mr. Zeisset asked the subgroup members for their suggestions about how the Bureau could make use of the feedback it received every day from calls taken by its customer service office. Ms. Ashcraft said that many companies have their customer service representatives note each call, and these notes are summarized, then reviewed by management. Dr. Etzel said another approach was to give incentive awards to the customer service representatives who first identify the problems. Ms. Stershic added that it was important for management to try to elicit these problems from its customer service representatives at daily debriefings.

Overview of the Census 2000 Dress Rehearsal (PAA/ASA)

Mr. Thompson noted that the budget agreement concluded by the Administration and the Congress in the fall of 1997 led to several modifications in the planning for and implementation of the dress rehearsal—

- At the Columbia, SC, site, the Bureau will no longer conduct a dress rehearsal based on using sampling for nonresponse follow-up, for following up vacant housing units reported by the post office, and for integrated coverage measurement (ICM). The dress rehearsal will include a post-enumeration survey which will be used to evaluate coverage but not to correct the population counts.

- The Bureau will conduct a thorough evaluation of the accuracy of both sampling and nonsampling methods.

- The data to be published for both the Sacramento, CA, and Menominee, WI, sites will also be displayed without the results of sampling and estimation. These data should be available on the Internet and on compact disk (CD–ROM).

- The Bureau will produce a midterm report on the dress rehearsal in the summer of 1998, the data from the evaluations by December 31, 1998, and a written report on the evaluations by January 31, 1999. These materials will serve as background for the decision, to be made in February, on whether Census 2000 will or will not use sampling.

A report is also being prepared describing the strategy for developing a plan for a nonsampling census. The report will be delivered to Congressman Rodgers (R, KY) by April 27, 1998. In addition, the agency's fiscal year (FY) 1998 budget legislation mandated the establishment of a Census Monitoring Board, consisting of eight members (four selected by the Administration and four by the Congress), whose charge is to review the planning and implementation of Census 2000.

When the Bureau last met with this Committee in October 1997, the agency did not have a budget and was operating under a "continuing resolution." Because of the agency's inability to award contracts (largely for questionnaire printing and data-capture equipment) while being funded by a continuing resolution, the dress rehearsal had to be postponed for two weeks.

Turning to the status of the dress rehearsal, Mr. Thompson said that the Bureau has completed the first phase of the local update of census addresses, and the U.S. Postal Service checked the areas with city-style addresses in March 1998. Promotion and outreach activities for the dress rehearsal have begun, as has the evaluation of the effectiveness of these activities.

He pointed out that there were two types of questionnaire delivery in the dress rehearsal, both of which have been completed—

- Update list/leave, in which Bureau employees deliver questionnaires and update their address lists, in areas with non-city-style addresses

- Delivery of questionnaires to housing units by U.S. Postal Service personnel. Mail delivery also included an advance letter, a reminder card, and a blanket mailing of replacement questionnaires.

For Census 2000, the Bureau has suspended plans for a blanket mailing of replacement questionnaires, due largely to the possibility of the program generating negative publicity and not producing the higher response rates for which the Bureau had planned and to the potential risk of receiving a large number of duplicate responses. Agency staff will evaluate the effects of the blanket second mailing in the dress rehearsal. Mail response rates in the three dress rehearsal sites as of April 21 were as follows:

Site	**Mail response rate (in percent)**
Menominee County, WI	34.9
Sacramento, CA	44.9
Columbia, SC	44.5

By the end of the mailout/mailback period, the agency hopes for a 50–percent response rate in Sacramento and 55 percent in Columbia.

The early parts of the enumeration of special populations have been completed. Census data were collected from individuals in transient locations on the night of April 17. Enumerators conducted the service-based enumeration of those without fixed addresses between April 20 and April 22. Enumeration of group quarters and of the military base at Fort Jackson, SC has begun. Bureau staff have also distributed Be Counted forms to appropriate locations in the test sites.

Data capture (including imaging, scanning, and keying from image) went into full production in the Jeffersonville, IN facility on April 20.

Telephone calls for assistance in receiving or completing a census questionnaire have totaled 15,852 to date. In addition, Bureau offices received about 600 phone calls in Spanish and 38 calls in other languages. Over half of these were resolved by an interactive voice recognition (IVR) system rather than being referred to an operator. The agency is extremely interested in the IVR rates because of their implications for the cost of telephone assistance in Census 2000. To date, only a few hundred respondents have asked to give complete interviews to Bureau personnel over the telephone. It is also worth noting that the volume of phone calls increased sharply just after the mailout of the replacement questionnaires.

Turning to recruiting enumerators for nonresponse follow-up, Mr. Thompson reminded the audience of the importance and challenge this operation will represent in 2000. The following table shows the Bureau's hiring goals for each site and the number and percent of qualified applicants for those positions.

Site	**Hiring goal**	**Number of qualified applicants**	**Qualified applicants as percentage of goal**
Menominee, WI	200	179	89

Sacramento, CA	6,990	4,843	69
Columbia, SC	11,575	5,218	45

The first two sites are basically on schedule, but Columbia should also be at about 69 percent. The unemployment rate in Columbia is very low, and there is significant competition for temporary workers. At the Columbia site, the Bureau has taken several steps to attract more qualified workers—

- Raised enumerator pay by $1.00 per hour.
- Increased publicity.
- Will examine other strategies next week.

The agency will carefully analyze the rate at which workers accept temporary positions and remain on the job. Bureau assumptions include a 200-percent turnover rate for employees; lower turnover rates would reduce the need for such a large applicant pool.

The Bureau finished the first phase (independent address listing) of the integrated coverage measurement (ICM) program and is currently matching the housing unit part of the address list to the dress rehearsal address list. Nonresponse follow-up will begin on May 14 and ICM personal visit interviews will start on May 29 in the Menominee and Sacramento sites. Post-enumeration survey personal visit interviews in the Columbia site will begin on June 12. These will start on a flow basis as each tract finishes its initial data-collection operations. The agency estimates that ICM interviewing will take about 8 weeks.

The local update of census addresses for Census 2000 has been under way for some time. The Bureau has received between 12,000 and 13,000 responses to its request for state, local, and tribal governments to participate in this program; about 9,000 of these governments have agreed to participate in the program.

Finally, on April 1, 1998, the agency submitted the actual questions to be used on Census 2000 questionnaires to the Congress.

Sampling and Estimation in Census 2000 and the Dress Rehearsal (PAA/ASA)

Dr. Hogan's presentation focused on the two Census 2000 Dress Rehearsal sites in which sampling and estimation will be used—Menominee County, WI, and Sacramento, CA. He summarized the four major sampling and estimation procedures the Bureau plans to use in Census 2000—

- Nonresponse follow-up.
- Undeliverable as addressed vacant follow-up.
- Integrated coverage measurement.
- Service-based enumeration.

The nonresponse follow-up sampling plan aims to achieve a 90-percent response rate in each census tract. For example, if a particular tract had a 60-percent response rate, and therefore a 40-percent nonresponse rate, the nonrespondents would be sampled at a 3-in-4 rate to bring the response rate in the entire tract to 90 percent.

In cases where the U.S. Postal Service returns a questionnaire to a census office indicating that the address used was vacant, the Bureau originally planned to conduct a 1-in-10 sample to verify the status of these housing units. This Advisory Committee, among other stakeholders, suggested the Bureau use a higher rate. Upon reconsideration, the agency has increased the sample percentage to 30 to assure that the Postal Service was not making systematic mistakes.

The goals of integrated coverage measurement (ICM) are to measure state populations directly and to remove the differential undercount. The agency hopes to obtain a coefficient of variation (CV) of 5 percent or less for each state. The components of the ICM are sampling, listing of housing units, housing-unit matching (a new operation for the dress rehearsal and Census 2000), interviewing (some will be telephoned), person matching, follow-up, and final matching and coding.

The sampling plan calls for a 750,000 housing-unit sample for the ICM in Census 2000. The samples will be state based and will permit direct estimates of state populations. Within states, the plan anticipates proportional allocation of the sample.

The steps involved in ICM estimation include missing data adjustment, dual system estimation, carrying the estimates down to small areas, and person imputation. Post stratification categories include tenure, race, ethnicity, age, and sex. One issue that was not present in 1990 is the Bureau's plan for handling multiracial responses. There are 126 combinations of race and Hispanic origin. The challenge will be to create categories of sufficient size to allow for the computation of the dual system estimate. For the dress rehearsal, the Bureau plans to place anyone who checks White and any other race in a "minority" poststratum. This approach should help the agency measure the differential undercount. When a respondent checks two race groups other than White, the plan calls for placing the respondent in the larger minority group poststratum. The Bureau has very little experience with multi-race responses and will evaluate the dress rehearsal data very carefully to ascertain how effective these procedures were.

As a result of previous Committee recommendations, the Bureau is looking at performing ICM raking within states in 2000 and within sites in the dress rehearsal. The agency will do two-way raking, with owner-renter as one dimension and age-race-sex as the other dimension.

He announced that the ICM will not be combined with demographic analysis estimates in Census 2000, largely because the Bureau could not identify one methodology that was clearly superior to the various alternatives. In addition, the agency will not produce an official household data file as a Census 2000 data product. Given the status of the research on household reconstruction, the Bureau determined that it could not produce a professional, high-quality product and therefore would not release this file as an official data product. A research data file that will incorporate what is learned between now and Census 2000 will be released, but it will not be an official data product.

Dr. Klerman of the Population Association of America (PAA) subgroup noted that many of the policy decisions concerning Census 2000 will be made by political appointees rather than by the career officials that attend these Advisory Committee meetings. The Committee members tend to focus on technical issues, although they are aware of the political context in which those issues are played out. Among knowledgeable professionals there is a consensus that the census-taking methods used in 1990 did not work then, will not work in 2000, and that sampling would be an appropriate alternative. The 1990 methodology missed large numbers of people and missed them disproportionately. He expressed concern about any proposal to return to the procedures used in 1990. The correct question to ask is not whether sampling and estimation procedures are perfect; rather it is whether these procedures are better than those used in 1990. There is almost universal agreement among demographers and statisticians that it is possible to design sampling procedures that will perform better than the procedures used in 1990.

No matter what decisions are made at this meeting, debates over data collection will continue after the meeting is over. The Congress and the courts will decide what the final numbers will be.

He urged the Bureau to make every effort to preserve integrated coverage measurement or the post-enumeration survey.

He expressed some frustration that despite the consensus among experienced professionals about the efficacy of taking a census involving sampling in 2000, the existence of this consensus does not appear to have reached the general public. At least part of the responsibility for failing to convey the consensus to the public belongs to the media.

On the more technical level, he expressed four concerns—

- The feasibility of the Bureau's plans for Census 2000, particularly the effort to hire large numbers of temporary workers (plans that would be exacerbated if the Congress requires the Bureau to conduct a full-scale nonresponse follow-up operation). In the event of a shortage of labor or funding, he urged the Bureau to consider reducing the size of the ICM. He added that bias due to incorrect enumeration would be much greater than sampling variance.

- The complexity of the estimation procedures. The statistical procedures proposed for the 1990 census were so complex that the Bureau had trouble describing them correctly and then implementing them. He argued that Census 2000 should favor the simplest procedures that will produce good estimates. Simpler procedures are easier to implement, explain, and document. A focus on simplicity would also help minimize the criticism that the Bureau's plan is so complicated that it would make it easier for politics to influence the results. Prespecification of a draft estimation algorithm would also help counter charges of potential political manipulation. The Bureau should consider sponsoring a workshop to be held in about 6 months, prior to which it should release to the public a draft algorithm, computer code, and dress rehearsal data. Following discussions with supporters and critics but well before census data collection begins, the agency should release the algorithm and computer code (subject to confidentiality restrictions) that will be used to process Census 2000.

- A more explicit schedule for key remaining elements of the planning for Census 2000. This would include timetables for receiving responses from the dress rehearsal sites and for the release of the Bureau's proposed ICM algorithm, the public response period, and the publication of the Bureau's final decision.

Dr. Tourangeau of the American Statistical Association (ASA) subgroup summarized the steps involved in sampling and estimation and expressed concern about the balance of complexity, accuracy, and timeliness that will be needed to complete this process in time to present state population figures to the President by December 31, 2000. He noted that the Bureau expects to complete the sampling and estimation phases of Census 2000 by December 1; during the 1990 census, the agency was able to complete all the necessary work, but only by July 15, 1991.

Like Dr. Klerman, Dr. Tourangeau stressed the importance of balancing complexity with accuracy and timeliness. Wherever possible, simplicity is preferable with regard to operational feasibility, ease of explanation, and other reasons. He proposed several principles the Bureau might consider for reducing risk in Census 2000—

- Consider dropping steps, such as imputation of nonresponding nonsample cases. He suggested substituting weighted estimation and controlled rounding. Perhaps iterative proportional fitting (raking) could also be dropped from the ICM process.

- Do as much as possible ahead of time. For example, many of the ICM sampling activities could be done in advance.

- Simplify as much as possible. He wondered if it would be possible to simplify the treatment of inmovers and outmovers.

- Build in checks to assure that what is done is done correctly. He felt sure the Bureau had worked on this and would like the agency to explain steps it has taken.

- Increase speed of operations. To meet the shorter schedule envisioned by the Census 2000 plan, the Bureau will be using improved technology for matching and will incorporate telephone interviews into the ICM.

On a more technical level, he wondered if, in determining the size of the nonresponse follow-up sample, it would be possible to remove the undeliverable-as-addressed vacants from both the numerator and denominator of the response rate. Using the Sacramento, CA, data, the Bureau could check the impact of this on overall sample size and on allocation by area.

He noted that the Bureau's plan to impute nonsample returns included the advantages of using systematic hot-deck imputation instead of weighting. He wondered why weighting was described in the background paper ("Sampling and Estimation in Census 2000 and the Dress Rehearsal") as "nearly unbiased." He expressed concern about the Bureau's plan to use controlled rounding to assure the release of block-level data in whole numbers. Another procedure that concerned him was the treatment of late mail returns, particularly the setting of a "drop dead" date after which no further mail returns would be accepted. He also wondered which imputations would be excluded from the ICM matching procedure. He thought the Bureau should be prepared to deal with a substantially larger number of erroneous

enumerations in 2000 than in 1990 because of the increased outreach effort, more effective advertising, and the "Be Counted" forms and wondered about the potential impact of this on the dual system estimates. Finally, he worried about the tightness of the time schedule for Census 2000 and the possible impact of delays in one procedure on subsequent operations.

Dr. Myers of the PAA subgroup noted that the background paper indicated the Bureau does not plan to release a household data file. This would mean that there will be a population count and a housing-unit count but they will not match each other. The integration of census operations, sampling and estimation, and data products in Census 2000 does not appear to be as close as he thought they were going to be. Dr. Hogan replied that the data products from Census 2000 will probably look more like the 1990 products than Dr. Myers will like. The Bureau had hoped to resolve this issue more favorably but was unable to do so in the time available. Bureau statisticians are working with demographers in the Population Division and with data users to determine the most useful ways to display the data. Ms. Schneider added that this is an extremely complex topic and suggested that the Committee might want a fuller discussion at a subsequent meeting or in a conference telephone call.

Dr. Stolzenberg of the PAA subgroup said that there were enemies of an accurate count and that Bureau efforts to be clear and correct in its decision making may not be sufficient. With regard to the issue of contracting out key components of census operations, the Bureau will face the challenge of specifying in detail what is expected of the contractors while allowing sufficient flexibility to deal with complications as they arise.

Ms. Becker of the PAA subgroup noted that in 1990 the census was prestratified by geography. She wondered if a similar process was planned for Census 2000. Dr. Hogan noted that although this was not incorporated in the dress rehearsal, it is under consideration for Census 2000. Geography is central to the sampling and estimation procedures for Census 2000 in that they will be done on a state-by-state basis with proportional allocation within states. Mr. Thompson added that proportional allocation has an implied stratification by geography. Ms. Becker pointed out that in 1990 the Bureau calculated undercount rates for Detroit and Chicago by stratum. Dr. Hogan replied that one 1990 poststratum did contain both Detroit and Chicago. With the state-based estimates planned for 2000, these two cities would be separated. One of the conclusions drawn from the 1990 census was the importance of tenure, which accounted for much of the distinction between central cities and other places.

Dr. Stokes of the ASA subgroup spoke in favor of complexity, noting that the Census Plus methodology[1] was simpler than the alternatives but that it did not work in practice. This should be kept in mind in the context of the earlier discussion about making things easier to explain to a broad audience. On the issue of a smaller sample size for ICM, the larger ICM sample proposed for Census 2000 was intended to allow poststratification within states. In this context, reducing the ICM sample is likely to increase bias as well as variance.

---

[1](Editor's note.) An alternative to dual system estimation, Census Plus involved taking a post-census survey using intensive enumeration techniques (e.g., administrative records, ethnographic enumeration, and highly trained and experienced enumerators) that were too costly or difficult to be used in a complete census. Statisticians would calculate the ratios of Census Plus enumerations to the regular census and use the ratios to estimate the total population.

Dr. Bell of the ASA subgroup agreed with Dr. Stokes' comment on ICM sample size. He supported the Bureau's plan for iterative proportional fitting (raking) but asked the agency to provide its plan for evaluating the dimensions for raking. He approved of the Bureau's plans for dealing with late census returns and with limiting any block cluster in the ICM to 80 housing units.

Dr. O'Hare of the ASA subgroup noted the importance of possible labor shortages in 2000. Dr. Klerman agreed and asked about the size of the work force the Bureau expected to recruit in 2000. Mr. Thompson replied that the Bureau wants to identify about 2.7 million qualified applicants to work in Census 2000, from which the agency expects to train approximately 500,000. The Bureau plans to set pay rates at levels sufficient to attract applicants who are already employed to supplement their income; the entire effort is not aimed solely at hiring the unemployed.

Overview of Indicators of Innovation and Technology (AEA)

Dr. Knickerbocker introduced Mr. Robert Shapiro, the Undersecretary for Economic Affairs at the Department of Commerce.

Mr. Shapiro expressed his own and the Department of Commerce's appreciation for the efforts of the members in helping the Census Bureau design and carry out the most accurate and sound census possible. He noted that the Census Bureau is currently in the midst of a deep political controversy. The position of the Bureau and the Department of Commerce is that their most important objective is the development and application of professional protocols and measures for the census. He emphasized the importance of defending the statistical integrity of the census process and that his primary mission as Undersecretary is to protect the integrity of the census design process against any political interference. Politicians should not be involved in any questions of statistical or systems design; that is work for professionals. The census design should be dictated by the state of the art in statistical design and method, not by political considerations.

In reply to a question by Dr. Willis, Mr. Shapiro said that the Administration agreed last November to expedited judicial review of the Constitutionality and legality of using statistical sampling as part of the census design. He noted that he had recently reviewed the memoranda prepared by the Department of Justice under President Bush's Administration defending the constitutional mandate that the census be carried out in whatever form will produce the most accurate result. The Congress has the constitutional authority to decide how the census should be done, and the Census Act delegates that authority to the Secretary of Commerce, who has delegated the authority to the Director of the Census Bureau. As a matter of constitutional law, the Department of Justice is compelled to make the argument that those who are challenging the constitutionality of sampling have no standing to challenge. The Supreme Court does not render advisory opinions, hence persons cannot sue over harm done by the census 2 years before the census in question has taken place. If the Supreme Court feels there is justification for addressing the merit of the issue, as distinct from the procedural issue of standing, the Administration is confident that the Court will uphold the constitutionality of the census design.

In response to a question by Dr. Scherer, Mr. Shapiro said there are two suits, in two different district courts, but they will be consolidated when the case is carried forward. He

suggested that the Supreme Court will rule on the questions involved either near the end of this year or early in 1999.

Dr. Knickerbocker noted that three members of the American Economics Association (AEA) subgroup are absent, Drs. Ernst Berndt, Ariel Pakes, and Lynn Brown. Dr. Rebecca Maynard of the University of Pennsylvania is a new member attending her first meeting. He said the discussion will begin with Dr. John Haltiwanger, the Chief Economist of the Census Bureau, and Dr. Ronald Cooper, a consulting economist. Later in the afternoon, the subgroup will discuss the National Science Foundation's (NSF's) research and development (R&D) survey, followed by a panel discussion on measuring and collecting data on research and development.

Dr. Haltiwanger said this session is intended to be a "springboard" for the later discussions. These sessions are intended to respond to recommendations from the AEA subgroup that the Bureau take a look at the NSF's industrial R&D survey, and that the survey should be set into its broader context. Last year the NSF's Science Resource Studies Division sponsored a workshop that covered much of the ground the Bureau and the subgroup are on today. The agency does not want to rehash what was done last year, but to build on that work. He said Dr. Cooper, one of the editors of the report on that workshop, will give an overview of what was done at that workshop.

Dr. Cooper said he will run through some of the main points of the workshop, which was held at the National Academy of Sciences in February 1997. The objective of the workshop was to generate ideas to improve national statistical information on industrial innovation to help inform public policy debates. The discussion did not address the human resource metric for innovation, but other indicators. A separate workshop is being planned on developing human resource metrics to track industrial innovation.

The discussions began with a look at the demand for information on innovation, which included academic research (to develop an informational base to better understand the role of innovation in the economy and to be able to track important trends in innovation). The workshop also discussed specific areas of policy needs, such as the Congress' interest in understanding returns on public expenditures for industrial innovation, as well as the needs of specific agencies to measure the impact of innovation, as required by the Government Performance and Results Act. An official of the Federal Trade Commission discussed that agency's need to be able to measure innovative capacity and the effect of mergers and acquisitions on that capacity. Other agencies indicated they did ad hoc surveys themselves and need to have information on particular aspects of technology. Various private industries also participated in the workshop and showed great interest in ways to improve their ability to measure the value of their own research and development activity.

To meet these needs, the workshop included a session on developing an analytical framework. Several participants pointed out that a better conceptual framework of the process of innovation is needed to build a useful national data system. Much of the discussion concerned the factors involved in innovation, and it became clear that there is no solid or comprehensive model of the process of innovation. The fragmented nature of the innovation model, or the lack of any model at all, necessarily impacts the quality of the data available.

Dr. Cooper said a presentation by Adam Jaffe was particular useful; he used a chart that showed inputs and outputs that applied to innovation. This can be much further developed; e.g., many of the inputs are not simply investments; many are conditions of the firm and the competitive condition of the industry involved.

Dr. Cooper outlined some of the recommendations made at the workshop for improving the data. On the input side, participants stressed the need to improve research and development information, primarily from the Bureau's RD-1 Survey of Industrial Research and Development. The changes suggested were to use the business unit rather than the firm as the survey unit, expand the coverage, and make response mandatory (possibly reducing overall respondent burden by reducing the number of items on the questionnaire). There are also nonresearch and development costs associated with innovation, and there have been efforts to collect data on these costs in other countries, where surveys have asked firms for the costs associated with design, training, and marketing that go into the process of innovation. These are good indicators, and it would be useful to be able to track them. Other inputs to research and development could include management of research and development, the way the firm appropriates returns to innovation, and geographic location. He noted that geographic location came up repeatedly as a factor in innovation spillovers, with "Silicon Valley" cited as a prime example.

Output data needs included patent data (needed to understand the use of patents versus other means of protecting intellectual properties), patent citation data, and technology use data. More work also was indicated on productivity and measures of economic "spillovers." Physical location also came up as part of the need to link innovation to economic development. There was a presentation at the workshop on using financial market indicators to measure the value of research and development activities, and participants thought this kind of research should be encouraged.

Dr. Cooper said that the keys to improving existing data on innovation are (1) linking innovation to relevant economic data bases, (2) collecting research and development and other data at the business use level, (3) increasing coverage (of research and development surveys), (4) encouraging industry participation and cooperation, (5) assuring the presence of geographic identifiers, (6) developing new data items in a national survey or series of surveys conducted on a regular basis, (7) longitudinal data continuity, with room in the process to explore new issues as needed, and (8) clarifying the role of Federal agencies versus private agencies (e.g., deciding which data could be collected best by the Federal Government and which could be better collected by private agencies or by public-private partnerships).

A final recommendation from the workshop was that a consultative body should be formed to provide the various Federal agencies involved with expert advice on these issues.

Dr. Haltiwanger said there are several questions and issues that the Bureau hopes to cover this afternoon. Together with the material covered by Dr. Cooper, Dr. Haltiwanger said other points for consideration include, what is meant by core data that the Federal agencies should be involved with, and how might they best collect that data. Three related questions about these points are "on the table"—

- What should the Bureau do about the existing surveys, such as the RD-1?

- Is it possible to satisfy some of these data needs with special modules that might be added to one or more of the Bureau's existing surveys?

- Is a new technology or innovation survey needed? If so, how does the Bureau assemble the sponsors needed to obtain the resources required to do the new survey?

In response to a question by Dr. Betancourt, Dr. Knickerbocker said the Bureau has no preference, at least officially. He pointed out, however, that there is a fourth alternative to the three listed by Dr. Haltiwanger: Can the Bureau take existing data sets on research and development and, by linking them in new ways or to other databases, obtain new insights or shed new light on public policy questions? He noted that the NSF is the official sponsor of the RD-1 survey, and is, in effect, the steward of research and development statistics. The Census Bureau mounts over 100 surveys every year and carries out the economic census. It is quite feasible that the Bureau could add special modules to some of those surveys or to the census to try to collect additional information (he noted that it is not too early to start thinking about the questionnaires for the 2002 Economic Census, since work on the content will start in the next year or so). He said he was interested whether the subgroup could suggest the one, two, or three more facts that, if available, might significantly improve users' ability to think about or use research and development data. Once that is known, the Bureau can consider how to incorporate them into the collection activities.

Panel Discussion: The National Science Foundation Research and Development Survey

Ms. Champion said the National Science Foundation's (NSF's) Research and Development (R&D) survey has been conducted since 1957 to measure R&D spending in American private industry. The key data variables associated with the survey are total amount of spending, amount of Federal funding and the Federal agencies providing this funding, types of research being conducted, fields of research, product areas affected by the research, the number of scientists and/or engineers involved, and the states in which the research was conducted.

Until the 1980s, most R&D had been conducted in a small number of large manufacturing companies. In the 1970s, a sample of 10,000 to 12,000 manufacturing companies was conducted, of which approximately 3,000 indicated they performed R&D. About 1,500 of these companies accounted for 95 percent of the R&D activity, and only 200 accounted for 85 percent of the activity.

Between survey years, data were collected from these 3,000 companies, with particular attention paid to the companies that represented the majority of the R&D activity in the United States. This distribution remained relatively unchanged until a period of mergers, divestitures, etc. in the 1980s. During this time, many companies stopped replying to the survey, assuming they were included with their new parent companies.

While companies were being lost from the sample, there was also an upsurge in the number of smaller companies conducting R&D in the computer and biotechnology industries. Many larger companies began spinning-off their R&D units—forming smaller independent companies. As a result of these changes, the Bureau designed a new annual survey in 1992 that allowed for the redefinition of the sample to better reflect companies' distributions and

characteristics. The Bureau also opened the survey to non-manufacturing industries which had previously been assumed to have very little R&D activity.

Several refinements were made to allocate the sample better. More information has been collected on the active industries. This has helped the Census Bureau methodology staff design a sample that targets companies performing R&D.

Currently there are about 25,000 companies in the sample each year. Data from the last survey showed the number of companies reporting R&D activities had grown by approximately 2,000. Of these companies, approximately 200 were conducting 75 percent of the R&D activities in the United States.

The Census Bureau tries to anticipate potential problems, such as companies "disappearing" through mergers and acquisitions, and the survey asks respondents to indicate any changes in each company's status. Respondents are also offered an electronic reporting option, which uses a diskette with built-in edits. A more systematic approach is being developed to compare this information to the data in the Security and Exchange Commission's annual reports. The content of the NSF's survey itself has remained unchanged. Additions have been suggested, but these are still being studied.

Dr. Scherer said data subdivided geographically below the state level are needed. Proximate entities tend to derive much more benefit than distant entities. Collecting data by Standard Metropolitan Statistical Area should be studied, because, for example, Wilmington is much more likely to receive spill-overs from Philadelphia than Pittsburgh. This is the logic underlying the desire for greater geographic segmentation.

Regarding the breakdown of total company-financed R&D by company size, he noted that by using the Gross Domestic Product deflator, there was a sharp decline in R&D spending in companies with more than 5,000 employees in the first tier of the breakdown, a decline in the second tier, and a modest increase in the third tier. The smallest firms reported a huge increase in R&D. He asked if this was a result of broader sample coverage.

Mr. Long commented that the number of companies surveyed in 1992 increased from 3,000 to 5,000, due to the addition of more small companies to the sample. He did not believe tabulations were performed in 1991 and 1992, when the increase took place, to determine what effect the 2,000 additional firms had on the data.

Dr. Scherer suggested it would be interesting to have an analysis of three effects—a constant panel without changes for small firms, the introduction of more small firms into the sample, and the introduction of particularly high-technology firms.

Mr. Long said that there were, particularly in the Interstate 270 corridor in Maryland and the Route 128 corridor in Massachusetts, hundreds of new biotechnology firms in 1993 that had not existed in 1990. This represents not only a change in sample design, but also a change in corporate reality.

Mr. Jankowski noted that a few special tabulations were performed to help understand the new sampling structure. These statistics were reported in 1992 at the aggregate level. When finer details are required, problems with industry shifting and new firms and industries

are encountered. The 1992 frame had not been changed since 1987. It should be expected that there would be an increase in the number of companies reporting following this 5-year period.

Dr. Scherer suggested tracking Initial Public Offerings (IPOs), since many represent small, high-technology firms. Mr. Jankowski said IPOs should not be a prerequisite for being a part of the sampling frame. Dr. Cohen added that patent data could also provide information on up-and-coming small high-technology firms.

Mr. Jarmon said there are many issues that can only be addressed with microdata. The Center for Economic Studies provides a place for researchers to access microdata files in Washington DC, Boston, and Pittsburgh. The Census Bureau data can be linked to other business data dating back to 1972.

There are several drawbacks to the microdata files. Most of the data collected are at the establishment level. It is difficult to allocate R&D expenditures across a company's various activities. The R&D survey is voluntary, which results in "spotty" coverage. Finally, the data are not collected and processed with microdata research in mind. Microdata researchers often find the cell sizes are too small to link to other data.

The survey deficiencies can be improved for microdata research if a better way was found to allocate R&D expenditures within a company's activities. This could be done by collecting data at the line of business level, improving geographic detail, making some of the questions mandatory, and focusing more on the purpose of the R&D.

Dr. Scherer asked what the quality was of the applied research and development breakdowns. Mr. Jankowski replied that nonresponse was approximately 50 percent [for applied R&D by product class].

In response to Dr. Scherer's question, Ms. Champion said a follow-up survey was not conducted following the revision of the survey's instructions.

Dr. Gort agreed that it is important to have more information on R&D expenditures when analyzing productivity and the effects of R&D. He believed product breakdowns are needed. Since response rates varied enormously, one could concentrate on sectors having the most reliable information. In addition, he noted that the scientific community has not agreed upon the distinction between basic and applied research. This question will generate meaningless answers unless some operational definitions are decided.

Following a discussion of two papers "R&D: Foundation for Innovation" and "R&D and Innovation Statistics" by Mr. Jankowski, Dr. Betancourt asked if it was possible to determine a company's salary and capital expenditures from the data collected from the R&D survey. Mr. Jankowski said the survey offers microdata researchers a starting point. The specific process for determining the expenditures on capital and salaries is complex.

Mr. Long discussed his paper, "Differences in Reported R&D Data on the NSF/Census RD-1 Form and the SEC 10-K Form: A Micro-data Investigation."

Demonstration of the Latest DADS Prototype (AMA)

Ms. Rowland distributed a handout describing one of the Bureau's Statistical Research Division-sponsored presentations regarding expectations for electronic technology developments between now and 2010. The presentation was related to preplanning activities for Census 2010.

Ms. Rowland explained that today's demonstration featured the Bureau's Data Access and Dissemination System (DADS) 1997, known as DADS97. The Bureau developed this prototype for two reasons—(1) as a proof of concept for a request for proposal for a contractor to actually develop the system and (2) to use it for beta testing, i.e., to interact with the Bureau's users to see what they think of a system like this. The DADS97 has a tool-based approach. She added that International Business Machines (IBM) was awarded the contract to build the DADS for Census 2000 in conjunction with its subcontractors, Oracle and Environmental Systems Research, Inc. (ESRI).

Ms. Moore gave a demonstration of the DADS97 prototype and its tools, such as the guide, browse, search, view a table, create a query, metadata, and mapping tools. She pointed out that the DADS97 used JAVA, not HTML, as a "stop" or "back" button was not included as part of the prototype. As a result, if one started a query, it had to be completed. She demonstrated how DADS97 can select and analyze variables and customize tables.

Dr. Spiro asked if the need for Adobe Acrobat software with the DADS97 would be inconvenient for data users. Ms. Moore said people could download the software easily from the Internet without even registering for it. Ms. Moore also noted that DADS97 has access to economic census and survey data and the 1990 Census Public Use Microdata Samples (1-percent) that users can summarize, cross tabulate, etc.

Dr. Etzel asked if IBM would be involved in the distribution of DADS in the future. Mr. Wynegar said DADS was just a tool or a delivery mechanism, not a product for sale.

Dr. Etzel asked if users could determine how current the data DADS were. Ms. Moore said the users would find that information in the metadata in the DADS.

Ms. Stershic (Chairwoman of the Board of the American Marketing Association) asked if the DADS97 were available to the public. Mr. Wynegar said it was not, but the final version would be available on the Internet; however, there might be some kind of charge for that. Ms. Moore noted that the Bureau beta tested DADS97, for which 159 users signed up on the Internet, 77-percent of whom logged in, 24-percent utilized the feedback mechanism, 21-percent completed an online evaluation form, and 9-percent sent their feedback via electronic mail. Among the participants, 45-percent were from government agencies, 25-percent from educational institutions, 8 percent from libraries, and less than 10-percent from the media, business, etc. Also, 41 of the Bureau's State data centers, one business information center, and two census information centers participated in the beta testing.

Ms. Moore said the agency learned from the beta testing that the interface was too difficult for novice users. Many users did not understand the concept of a tool-based system, performance was slow, the system was not user-friendly (i.e., a stop or cancel button was not available), and there was not enough system feedback provided to the user while the system was processing.

Mr. Adams asked what percentage of the participants were novice users. Ms. Moore did not know the percentage. Mr. Adams also asked which tools the Bureau thought would be used most. Ms. Moore said table-building tools would be the most used; however, novice users probably would use the browse or search tools.

Ms. Moore noted that, for DADS98, the Bureau was focusing on an interface for novice users designed from a user's perspective and was eliminating the requirement of JAVA Aplets software. DADS98 will be released in January 1999 and will give users access to the 1990 census, 1997 Economic Census, and 1996 and 1997 American Community Survey data products; a second release will take place in March 1999, with access to the Census 2000 Dress Rehearsal products.

Mr. Adams complemented the Bureau for considering both the technical and the users' perspectives in developing the DADS.

How Do We Evaluate the Dress Rehearsal and Census 2000 (ASA/PAA)

Ms. Killion said that the dress rehearsal's objectives are to demonstrate the feasibility of integrating the methods that have been tested separately in the Census 2000 testing cycle and to test components of a nonsampling census. Normally, the Bureau does not test procedures in a dress rehearsal, but last fall's agreement between the Administration and the Congress on the agency's budget for fiscal year (FY) 1998 made it necessary to test components of a nonsampling census, particularly in the Columbia, SC, site, where integrated coverage measurement (ICM) will not be used.

The objectives of the dress rehearsal evaluation program are to validate the plans for Census 2000, to measure coverage of people and housing units, and to evaluate the completeness and quality of the data. The evaluation program will produce three main products—

- A mid-term status report, to be issued in late September 1998.
- Quality assurance reports will be issued as they are completed, to be supplemented in September by an overall quality assurance report.
- A report that consolidates all the key evaluations from the dress rehearsal and is due at the end of January 1999.

The dress rehearsal will result in eight broad categories of evaluation reports—

- Questionnaire related evaluations. These will evaluate such issues as questionnaire content and how the folding of the questionnaire will affect completion.
- Master Address File (MAF) evaluations. One of the most important of these will assess the accuracy and completeness of the MAF.
- Coverage measurement evaluations.
- Coverage improvement evaluations.

- Promotion evaluations.
- Multiple response resolution evaluations.
- Nonresponse follow-up and field infrastructure evaluations.
- Technology related evaluations.

In each site, the Bureau will also measure the quality of the dress rehearsal. These standards are based on the actual results of the 1990 census—

- Completeness of the MAF.
- Questionnaire mailing strategy. The increase in response from follow-up mailings should be at least 6 percent.
- Paid advertising campaign. The increase in awareness of census activities should be at least 30 percent.
- Nonresponse follow-up. This operation needs to be completed on time. In addition, the Bureau will evaluate the number of final-attempt cases and try to hold proxy and unclassified cases to less than 0.6 percent.
- Multiple response resolution. After the completion of data collection and processing, the proportion of duplicate enumerations should not exceed 1.2 percent and missed housing units should not exceed 1.3 percent.
- ICM/post-enumeration survey (PES). The Bureau will evaluate whether the ICM was completed on time and whether the one-number census reduced the differential undercount. In the Columbia, SC, site, the issue is whether including the PES results would reduce the undercount. The agency will use demographic analysis benchmarks to evaluate ICM/PES results.

Dr. Stokes of the American Statistical Association (ASA) subgroup, noted that neither the background paper ("Overview of the Census 2000 Dress Rehearsal Evaluation Program") nor the presentation said anything about quality assurance plans for the ICM in Census 2000. For example, in 1990 about one-third of the PES households were reinterviewed. She was also interested in the Bureau's plans for evaluating "curbstoning" in Census 2000. The presentation did not cover evaluations of other procedures such as matching rates and success in locating movers.

The statistical community generally expects that sampling for nonresponse follow-up will improve the timeliness and reduce the cost of Census 2000 and that ICM will reduce the differential undercount. Theoretically, sampling for nonresponse may also improve data quality because enumerator quality would be enhanced by not having to hire as many and more time could be devoted to each interview, as needed. She suggested that during Census 2000, the Bureau collect some comparative data on different types of last resort information. Ideally, the Bureau should use sampling for nonresponse in some randomized blocks and complete

nonresponse follow-up in others. Then researchers could compare the results. Since this is impossible, she turned to the ICM blocks for another type of test. The current plan calls for complete nonresponse follow-up in ICM blocks and sampling for nonresponse in non-ICM blocks. It might be possible to compare data on the types of last resort information obtained in the ICM and non-ICM blocks.

With regard to enumerator quality, it would be desirable to have work history, performance on a screening test, and other relevant characteristics that could be linked to performance measures. This information would allow researchers to use models to predict enumerator performance.

She thought there might be a problem in 2000 if people who received the long form chose to complete and return Be Counted forms instead.

On the last issue, Ms. Killion replied that the Bureau is conducting a formal evaluation of this possibility during the dress rehearsal.

Ms. Becker of the Population Association of America (PAA) subgroup, pointed out that the reengineered program for creating the MAF and the local update of census addresses (LUCA) is not being tested in the dress rehearsal. She said she had read the evaluation of LUCA from Sacramento and would like to participate in drawing up the evaluation plan for the MAF and LUCA in Census 2000. As a result of reassessing the creation of the MAF by linking the 1990 census address list to the U.S. Postal Service's Delivery Sequence File, the Bureau has decided it must spend $110 million to update the MAF by canvassing all the city style addresses in the country. She felt the Bureau should use the block listings from the ICM and PES to evaluate the problems in the MAF during the dress rehearsal. However, she did not believe an evaluation of LUCA was possible because it was not sufficiently well established during the dress rehearsal.

She asked if interactive voice recognition (IVR) would be used to complete census forms or only to provide respondents with information. Ms. Schneider replied that it was only for answering respondents' questions. To complete a census interview over the telephone, a respondent will have to speak with a census employee. Ms. Becker said she hoped the Bureau would evaluate IVR in the dress rehearsal. She added that she hoped the evaluation program for the dress rehearsal would not be reduced in scope. Ms. Killion stated that some of the proposed evaluations had already been dropped.

Ms. Becker complained about the limited distribution of 1990 research and evaluation reports and said it was important to plan for disseminating this information early in the decade following the taking of the census.

Ms. Killion pointed out that most of the issues raised by the discussants were included in the evaluation program. For example, one evaluation of ICM will use computer-assisted personal interviewing to detect curbstoning. Another evaluation, conducted by an outside contractor, will look at Bureau dress rehearsal field operations, assess successes and failures, and make recommendations for change. A third evaluation will assess the agency's ability to fully staff each operation. While the Bureau has been asked to determine whether sampling for nonresponse follow-up increases data quality, she pointed out that the agency has been unable to devise a measure that is not confounded by other factors. As measured by completed

interviews, the quality of last resort (now called final attempt) cases is not much inferior to respondent-collected data. In many cases, the final attempt was actually the first attempt. She said the Bureau would study the comment on interviewer quality. She noted that appendix D of the background document listed all the dress rehearsal evaluations.

Most of the standards established for the dress rehearsal and for Census 2000 are quite high. Ms. Schneider added that the Bureau had attempted to look at the type of proxy respondents but that all that it could establish was whether or not the form was completed as a proxy or as a regular interview. The Bureau may have to reconsider the widespread distribution of blank questionnaires since it does not want to increase the number of completed duplicate forms.

Ms. Killion remarked that one of the MAF evaluations will assess coverage while the other examines the MAF-building process. One of the biggest problems in the dress rehearsal was that the MAF did not have the 100-percent canvass that is planned for Census 2000. To evaluate interactive voice recognition, the Bureau will be examining the number of calls and hangups, Plans for evaluating the dress rehearsal were not completed early enough and, as a result, some of the evaluations probably will not be completed. As soon as the Bureau knows which ones will not be implemented, it will inform Committee members. One of the potential new uses of administrative records is to match addresses from the MAF to the administrative data file and use this information to identify problems with the MAF.

Ms. Killion said that the Bureau has established eight quality review boards for these evaluations. Each board will consist of three or four people. Two or three members will come from non-decennial parts of the Bureau, and one will come from another statistical agency. These boards will track the evaluations through the methodological development and review process. The Bureau also has a set of study plans for these evaluations and will distribute part or all of the plans to Committee members.

Dr. Voss (PAA) asked if the 30-percent increase in awareness the Bureau plans to use to define success in the dress rehearsal might be too low. While 30 percent was comparable to 1980 and 1990, one might expect paid advertising to do a better job than the pro bono campaigns of earlier censuses. Ms. Killion agreed that it was reasonable to expect more than 30 percent but said there was no way to estimate how much more. The standard that will be used is derived from the 1980 and 1990 advertising campaigns. Dr. Voss asked about the role of demographic analysis in the estimates program. Ms. Killion replied that data from demographic analysis will not be incorporated into the final, one-number population estimate. However, demographic analysis will be a key evaluation tool, as it has been in previous censuses. Ms. Schneider added that demographic analysis has generally produced higher population counts for traditionally undercounted groups than dual system estimation. The Bureau investigated the possibility of incorporating the results of demographic analysis into the final population estimates but decided to maintain it as an evaluation tool.

Dr. Bell (ASA) asked about methods of evaluating the differential undercount in the dress rehearsal sites in addition to demographic analysis. Ms. Killion mentioned the post-enumeration survey in the Columbia, SC, site and added that the Bureau is developing other evaluation tools, such as logistic regression, for use in both sites.

Dr. Bell suggested the Bureau improve its cost modeling for the 2010 census. Currently, the model does not seem to have the sensitivity to do "what if" calculations.

Panel Discussion: Where Should We Go From Here? (AEA)

Dr. Knickerbocker moderated the panel discussion, which included presentations by six panelists followed by questions and answers from others in attendance. The presentations were a continuation of an earlier panel discussion regarding the Survey of Industrial Research and Development sponsored by the National Science Foundation and conducted by the Census Bureau.

The first panelist, Mr. Wyndrum of the American Telephone and Telegraph Company, discussed four criteria, in order of importance, that major corporations use to set their budgets for research and development (R&D)—

- The current level of spending on R&D usually determines the following year's budget, but with rapidly changing technology, this criterion is becoming less relevant.

- Listing the most important research requirement, followed by the other research needs according to their importance, enables corporations to determine their budgets.

- Sales, general, and administrative costs also determine the research budget; total R&D is allotted a few percentage points of sales, and pure research a few tenths of a percentage point.

- Competitive analogs taken from unbiased surveys, such as the National Science Foundation and Center for Innovation Management Studies surveys, can play an important role in establishing corporate research budgets.

There are five critical research and development needs for better industry data–

- Expand the National Science Foundation/Center for Innovation Management Studies survey for better coverage of key industries, such as information technology, electronics, software, and semiconductors.

- Data should provide insight into process as well as product R&D, and into service as well as product industries.

- Data on innovation and application R&D should not be used to split research as opposed to development; industry does not make this distinction.

- Central R&D should be distinguishable from R&D conducted at separate business units.

- For multi-industry firms, data are needed by business unit and by industry.

The next panelist, Dr. Cohen of Carnegie-Mellon University, gave a presentation entitled, "Thoughts on Technology and Innovation Statistics." R&D data can benefit public policy, aid management decision making, and provide a basic understanding of innovation,

technical advance, and productivity growth. The data should be collected at the 3- and 4-digit Standard Industrial Classification (SIC) level, by type of R&D, and by type of firm and performer. Information also should be collected on technology acquisition and its determinants. Some of the issues associated with data on R&D include—

- Data at the enterprise level are often inappropriate.
- There is no single questionnaire or sample frame that will satisfy all data-collection needs.
- Linkage at both the firm and business unit levels with other data sets is critical.
- Different questions should be asked at different times, e.g., R&D versus questions on determinants.
- Some variables, such as R&D, should be subject to a census, while others, such as determinants, are suitable for much smaller samples.
- Objective measures should be employed in any innovation survey.
- Variables to be collected should be recognized as "moving targets."

Another panelist, Dr. Griffith of the National Science Foundation (NSF), said her agency had various ways of monitoring research and innovation. The NSF's science and industry data were from the Census Bureau and its patenting data were from the Patent and Trademark Office. The agency conducted a pilot innovation study a few years ago and is planning a new one. Some of the current issues concerning the collection of R&D data include—

- The collaboration between industry and academia.
- The need for additional science and technology databases.
- Data quality and consistency, especially consistency within firms.
- A better understanding of the service sector.
- The globalization of R&D.
- The need for the NSF to work closely with the Bureau and its advisory committees.

Mr. Long added another issue, linkage at the microdata level. For example, linking patent data with R&D data was a tedious, time-consuming process that had to be done by hand. Once this work was completed, however, the data could be compiled rather quickly. The problem was making the staff available to do the linkage.

The next panelist, Mr. Landefeld of the Bureau of Economic Analysis (BEA), explained that his agency's decision not to fund all its satellite accounts, including its R&D satellite account, was entirely due to budgetary issues. The BEA's core account had first priority, but the agency is seeking additional funding for satellite accounts.

A number of issues exist regarding input data. The data need to be comprehensive and current. Multinational companies make it difficult to have a complete accounting system and often result in double counting. Consideration of royalty flows in conjunction with R&D expenditures would be useful. Other types of input issues not specific to Census Bureau data are improved depreciation schedules, the deflation of inputs, and quality adjustments.

As for output data, further work is needed on the quality adjustment of outputs, e.g., for high technology, banking outputs, and all the services that are difficult to measure. Another problematic area is non-market measures of output. Expanded satellite accounts, whether they involve household production, environmental accounts, or health sectors, will impact R&D data.

Dr. Scherer of the American Economic Association (AEA), asked if anyone, besides Edward Mansfield in the early 1980's, had tried to compile a R&D input price-index deflator. Mr. Jankowski of the NSF said he had applied Mansfield's work in his own research and extended or updated it, but he was not aware of anyone who had done similar work independently.

Another panelist, Mr. Dean of the Bureau of Labor Statistics (BLS), said his agency had made use of the RD-1 data from the Annual R&D Survey, which were central to his agency's work involving multi-factor productivity measurements. The BLS had undertaken a large study of the direct effects of R&D expenditures and benefits, and the effects on multi-factor productivity growth. The study included R&D company stocks, estimates of rates of return on these stocks, and estimates of R&D depreciation as a product of accumulating the stocks. This study was published in 1989, and its data were incorporated into the BLS' annual updated measure of multi-factor productivity in the non-farm business sector.

Some important issues and needs include—

- Most importantly, change the RD-1 form to include additional break-outs for detailed line of business.

- Provide greater geographic detail.

- For line-of-business data, questions on the RD-1 form should be mandatory.

The last panelist, Dr. Bean of Lehigh University, discussed the IRI/CIMS Annual R&D Survey. The survey distinguishes between process and product development, covers technical services, includes the inputs into R&D, and how these inputs are transformed into useful outputs. The survey also attempts to obtain output measures, such as the patent measure. Two other important information categories are the new-sales ratio and the cost-savings ratio, which deal with attributions of sales levels in products or services associated with R&D activities and cost savings. Data are collected at the firm, business segment, and laboratory levels. Data also are collected on central laboratory activities. Collecting data at these levels becomes problematic when linkage is required, so standard identifiers are necessary for firms, business segments, and laboratories.

In response to a question from Mr. Long, Dr. Bean said the data from the IRI/CIMS Annual R&D Survey were compared with 10-K data, which he referred to as the "due-diligence" process.

In response to a question from Dr. Scherer, Dr. Bean said initially about 75 of the 270 Industrial Research Institute's member companies participated in the survey. Now, there are about 110 of 275 companies participating. The Industrial Research Institute estimates that about 30 percent of industrial R&D expenditures is being accounted for by the participating firms. There are a number of reasons why firms do not participate in the survey. Some companies have a policy not to participate in any survey that is not mandatory. Other reasons are "being too busy," downsizing, or not having staff to complete the questionnaires. Sometimes key persons leave who had filled out the questionnaires in the past, and the companies no longer return the forms. Also, there is competition with other organizations surveying many of the same companies for what these companies perceive to be similar information.

In response to a question from Dr. Cohen, Dr. Bean said that, at the firm level, the chief technical officer was usually the person responsible for completing the questionnaire. In other cases, the person responsible had been designated as the Industrial Research Institute representative. At the segment level, the person responsible usually was designated by the Industrial Research Institute representative or the chief technical officer. The laboratory director was responsible at the laboratory level.

Dr. Knickerbocker asked why the survey did not include any data on the cost-saving ratio and why there was no information on processes showing how R&D activities resulted in cost savings. Dr. Bean said there were fewer than five companies reporting process data, which was surprising because process should be easier to report than product data. Mr. Merrill of the National Academy of Sciences asked how the utility of the survey's data was evaluated. What were the most useful items of information for the Industrial Research Institute's members? Dr. Bean said that his organization provided the Industrial Research Institute's members with the data, per their request, but that it did not attempt to obtain information on data applications. Ms. Grucza of the Industrial Research Institute said a committee composed of the Industrial Research Institute's member companies had put the questionnaire together according to their perceived data needs. Now, however, it was time to start recognizing general industry needs. Two focus groups representing the food and electronics industries were to be conducted in conjunction with the Industrial Research Institute's next annual meeting in 2 or 3 weeks to develop supplemental questions of interest to these industries.

Mr. Long asked if there would be any benefit gained if the Bureau were responsible for the survey. Ms. Grucza said the primary benefit would be the expanded sample size of the survey. Dr. Knickerbocker added that disclosure would be less problematic under Bureau sponsorship.

Mr. Jankowski asked if the users would have the same quality of service from the Bureau that they now receive from Dr. Bean. Ms. Grucza said people such as Dr. Bean are busy and are not always available, whereas Bureau staff generally would be available to handle most of the questions and requests.

Dr. Cohen said he had identified four areas of consensus from the preceding discussion–

- Line-of-business data were important to both public and private sector users, but mandatory reporting may be the only way to collect these data.

- Patenting data were problematic, since they were easy to misinterpret and misuse.

- International comparability of the data was becoming more important, but achieving this comparability might mean agreeing to report categories of little or no relevance to North American data users.

- The knowledge level of the respondents was critical to the quality of the survey data.

Dr. Scherer asked what percentage of the R&D expenditure data from the survey is derived from laboratories associated with operating establishments, and what percentage is derived from "self-standing" laboratories not associated with operating establishments. Dr. Dunne said that question was asked of the operating establishments on the Bureau's Survey of Manufacturing Technology and is available in that survey's publications. Dr. Scherer said he believed those publications showed that about 70 percent of the laboratories were associated with operating establishments. Mr. Jankowski asked if the percentage was for dollars spent or the number of operating establishments. Dr. Dunne responded that the percentage data were based on operating establishments. Dr. Bean said the IRI/CIMS Annual R&D Survey data were available for both the corporate and the business segment laboratories. Dr. Cohen said data from the latter survey showed–

- Thirty percent of the laboratories were at corporate headquarters.

- Forty percent were at production facilities.

- Thirty percent were "stand-alone" laboratories.

Dr. Gort (AEA) expressed two concerns–

- Most of the data discussed today concern inputs, not outputs. Patents are an inadequate proxy for innovation output, and econometric studies of the impact of R&D are not a substitute for basic data.

- Most of the data currently available are for R&D expenditures, and often show small companies with zero budgets for research, which can not be true.

Ms. Grucza said there is a need to look at what industry itself collects to meet its own needs. Then, Government data-gathering agencies can work with these industries to show how they can transform what they collect into what the Government is requesting.

Dr. Scherer made three general observations–

- Although many attending the meeting would like data on outputs as well as inputs, the relationship between these two variables is extremely noisy. Those in attendance may have to accept that obtaining output data can only be achieved through statistical analysis, not through specific questionnaire techniques.

- A consensus exists for obtaining R&D data by business segment, but doing so is a labor-intensive and costly process. This approach, however, is the only alternative for obtaining these data.

- Also costly and labor intensive, is linking input to output data. To link these data, it is necessary to have input data at the business segment level. In this way, output data can be inferred from patent data.

- Dr. Willis (AEA) said another issue was linking census to industry data. The advantage of census data was the agency's mandatory reporting requirement. The disadvantage was that making the data available to companies would be difficult because of disclosure issues. Dr. Haltiwanger said the Bureau was studying this issue. More information was needed concerning the cooperative ventures between the Bureau and industry that already exist, in addition to what more can be done.

How Do We Evaluate the Marketing Strategy for the Dress Rehearsal and Census 2000? (AMA, ASA, PAA)

Ms. Bates said the evaluation of the dress rehearsal marketing strategy would determine if awareness in South Carolina and Sacramento, California increased as a result of the Bureau's paid advertising campaign. To evaluate the impact of the advertising campaign on response rates and awareness, the Bureau is using a 2-wave survey research design, administered by a random-digit dial survey. The evaluation does not include the Menominee site because its population would be too small to conduct this type of research design.

The first wave of the evaluation, designed to provide a "clean" measure of awareness, began prior to the start of the advertising campaign. The questionnaire used during the first wave was similar to those used during awareness surveys in the 1980 and 1990 censuses. In addition to asking people basic demographic questions, respondents were asked to comment on their media habits, civic participation, attitudes about the census, and awareness of census activities. Those respondents who indicated some awareness of the census were asked to list the sources contributing to their awareness.

On April 24, 1998, the second wave of the survey was initiated to conduct 1,500 post-advertising campaign interviews. In addition to asking questions similar to those in the first wave, respondents were also asked to identify the media outlets that increased their awareness, recall any of the census advertising, and indicate if the advertising campaign had an impact on their response.

After the two waves are completed, the dress rehearsal evaluation will compare differences between the two sets of interviews for changes in awareness. The Bureau's goals for the campaign are to increase awareness by 30 percent and to determine if increased awareness of the census had an impact on response rates.

Ms. Bates said the evaluation had the following limitations:

- There were variables affecting awareness that could not be attributed to the advertising campaign, including the partnership program and awareness resulting from the receipt of a census questionnaire.

- Several problems, including nonresponse, prevented the collection of 1,000 surveys during the first wave of the evaluation.

- The dress rehearsal will not get the same exposure as a national census.

Dr. Bell of the American Statistical Association (ASA) subgroup said the dress rehearsal's two-site sample size could present problems. He believed it would be difficult to determine if an increase in response and/or awareness was a result of the advertising campaign or the questionnaire mailout. With only two sites providing data, it will be difficult to determine if one form of advertising was more effective. Additionally, biases in the data may occur because the sample size is not large enough to compensate for those people who were interested and responded to the census regardless of the advertising campaign.

Dr. Bell suggested that the Bureau ask itself how much it should spend on advertising. This will be difficult to decide following the evaluation. The Bureau needs to identify what public attitudes need to be changed, especially concerns about confidentiality. The public's perception of census confidentiality may dictate where advertising money needs to be spent.

Ms. Bates agreed that the evaluation will not determine which forms of advertising were most effective; however, the evaluation questionnaire did include questions asking respondents if they could recall census advertising. Response to these questions may suggest that certain forms of advertising were more effective than others. The Bureau will break down the data from the evaluation into subgroups to identify if particular audiences should be targeted with specific advertising.

Dr. Stolzenberg of the Population Association of America (PAA) subgroup said that the evaluation may not gain much, but it will not cost much, since the marketing has to be performed as part of the dress rehearsal operations.

Dr. Jacobsen (PAA) said that if the Bureau wanted to measure the effect of advertising on awareness, independent of the effect of the mailout of the questionnaires, a three-wave design should have been conducted. The three-wave design would have included interviews before any advertising took place, after the advertising had begun, but before the forms were mailed, and following the mailout of the questionnaires.

Ms. Bates said that the Bureau would have liked to conduct a three-wave evaluation, but timing and other factors prevented it. The second wave questionnaires did include follow-up questions that asked if the respondent had heard about the census only after receiving the census form.

In response to a question by Dr. Jacobsen, Ms. Bates said the questionnaire for the second wave included two additional questions designed to measure the "innovative and aggressive" advertising in South Carolina. The frequency of responses to these questions will be analyzed to determine the effectiveness of this non-traditional advertising.

Dr. Jacobsen said that the Bureau should expect to increase awareness more than 30 percent following a $100 million advertising campaign. However, the goal of a 30-percent

increase for the dress rehearsal may also be too high. The dress rehearsal is missing the national advertising that is present during a decennial census. She expressed concern that the Bureau did not have adequate reasons for establishing the 30-percent increase figure.

Considering that response rates to all types of surveys have been decreasing, the Bureau may be more interested in looking at rates of change rather than actual response rates from other national surveys. There are a number of companies that maintain statistics from mailout/mailback surveys. Although these may not be exactly comparable because of incentive offers associated with these private surveys, an average of response rates could be studied.

Dr. Jacobsen expressed concern over the legitimacy of the evaluation data, given the small sample size. Because a portion of the first wave interviews were conducted after the paid advertising had begun, there will be a lack of credibility and validity in the results. Ms. Bates said she had received the "clean and weighted" data set from the first wave on April 22, 1998. Prior to this, she had been looking at the data collected in the first wave, which included the late interviews. Although there have been few differences in the responses during the first wave, the differences they have seen may require that these interviews be excluded from the evaluation.

Ms. Ashcraft of the American Marketing Association (AMA) made the following observations:

- The inclusion of measurements at key points of the marketing hierarchy of effects model, which builds from awareness, attitudes, intent, and behaviors, was a good idea.
- Capturing both pre- and post-awareness was necessary because of "ghost awareness in the pre-wave."
- The two surveys should not be longitudinal.
- A separate sample is preferred in each case so as not to sensitize respondents the pre- and post-waves.
- It was good that the pre- and post-waves were weighted demographically making them comparable.
- The use of multi-variant statistics was a good idea.
- It would be interesting to see if one subgroup lagged behind, or if one responded in a particular way.
- It was a good idea to have a "regression" performed against the census questionnaires as the dependent variable, helping to identify the driving behavior.
- It was unfortunate that the evaluation did not achieve its respondent quota in the first wave.
- It was unfortunate that the dress rehearsal did not include a control city where there was no paid advertising.

- It would be difficult to separate awareness from the various forms of marketing. Traditionally, respondents are often unclear as to where they saw something. It may be effective to have respondents recall specific visual elements or copy points.

- It is important to differentiate between a 30-percent and a 30-percentage point increase in awareness. A 30-percentage point increase will depend on awareness during the first wave of the evaluation.

- It is best to use media weight awareness models that will connect media weight to awareness levels. This will be a good point of comparison throughout the evaluation.

- It will be difficult to compare the dress rehearsal to the census. It is difficult to compare a localized dress rehearsal to a national census. A national census' high degree of media coverage will increase national awareness. Additionally, the length of the Bureau's public awareness campaign will be much longer during the census than it was during the dress rehearsal. This will increase the chance that people will come into contact with census issues.

- Partnership and advertising activities are much cheaper and easier to coordinate on a localized level, like the dress rehearsal. On a national scale, it may not be financially feasible to conduct the same level of promotion.

- The advertising campaign can not change from the dress rehearsal to Census 2000, or the dress rehearsal will not be comparable. The advertising campaign needs to remain constant.

Dr. Meyer said the Bureau does plan to change the advertising campaign following the dress rehearsal. Many of the design elements currently being evaluated will not be used during the Census 2000 campaign. Several new elements, especially for targeted groups, will be added. Initially, the dress rehearsal was to test some themes and messages as part of a research and development program leading up to the Census 2000 campaign. The oversight committee recommended the campaign include an evaluation of the advertising campaign in the dress rehearsal. Had a full dress rehearsal of the advertising campaign been planned from the start, many things would have been done differently.

Ms. Ashcraft said that if revisions were made to the advertising campaign for Census 2000, some persuasive research showing why these changes were made should be included. Dr. Meyer said that the Bureau intended to have an extensive research program budgeted in Fiscal Year 1999 to evaluate and research advertising campaign changes.

In response to a question from Dr. Spiro (AMA), Dr. Meyer said a question was included in the first-wave survey that asked respondents how likely they would be to complete and return the dress rehearsal questionnaire.

Dr. Etzel (AMA) said that in the businesses test-marketing products will often remove the product from that market if there is local publicity. Publicity can adversely affect test-product research. The Bureau should remain cognizant of publicity that is outside the dress rehearsal's efforts.

Ms. Becker (PAA) asked if the first-wave response rates were available. She was concerned about the dress rehearsal's literal representation of the population and asked if there was any way to account for people without telephones or those who were unwilling to respond to the survey. Ms. Bates said the early data on response rates was not good. This was one of the reasons the first wave of the evaluation took longer than expected. A decision will need to be whether the late responses to the first wave will be included in the evaluation.

Ms. Becker asked if the results of the survey, including those items on confidentiality, would be made available. She would like to see accurate data as opposed to speculation about how the public is reacting to census confidentiality. Ms. Bates said there were a few questions addressing confidentiality.

Mr. Adams (AMA) said it would have been helpful if a control cell were included in the evaluation. He was concerned that the Bureau would attribute awareness to its own advertising campaign rather than to outside variables like the national and local newspapers.

He believed the analysis and objectives for the evaluation may be too biased toward studying awareness. Awareness is more of a commodity in the marketing sense than familiarity, commitment and intent. He was encouraged by the regression, but agreed with the other members that the lack of observation points may make this difficult. He hoped the dress rehearsal would provide some information on how strong an impediment confidentiality will be to response rates.

How Should the Census Bureau Price Data Products Through DADS (AMA, PAA)

Mr. Kavaliunas told the members of the American Marketing Association and Population Association of America subgroups that an interdivisional team was established in December 1997 to discuss census products and pricing. The interdivisional team reviewed the costs associated with the dissemination of census products, developed a pricing scheme for census products, and sought both internal and external discussion of the Bureau's pricing of products. The Bureau has focused its efforts upon the development of electronic delivery of census products. Coincidentally, the sale of printed census data had declined over the past decade according to the sales figures from the Government Printing Office.

The interdivisional team recommended the Bureau continue to offer basic information to users via the Internet at no charge. There would continue to be a charge associated with off-the-shelf products, downloads of large files, creating custom products, and any product and/or tabulation requiring Bureau staff time. The Bureau is working with the United States Department of the Treasury to be able to process credit card orders for products through the Internet. We expect this procedure to be in place this summer.

Mr. Kavaliunas told the two subgroups that the Bureau's pricing of data products must follow the Federal information policy found in the Office of Management and Budget's (OMBs) Circular A-130. This policy prohibits the Bureau from entering into restrictive or exclusive arrangements for its products and requires the Bureau to price its products so as to recover the cost of dissemination. The price of data products would be determined by the costs associated with preparing products for dissemination and actually providing these to data users. The cost of dissemination is estimated at $10 million per year for the years 2001 through 2005. This cost

reflects the number of full time employees involved in data dissemination (approximately 10 percent of the Bureau's workforce).

In addition to the sale of data products, the Bureau is studying several ideas to increase revenues, including the addition of basic Geographic Information System (GIS) capabilities within DADS and enhancing the Internet subscription service that is currently available. The Bureau also is considering high-speed and dedicated lines for frequent users and hosting data for other Federal statistical agencies.

Mr. Kavaliunas asked the subgroups for their comments on the pricing of data products and ideas for revenue-generating proposals.

Dr. Voss of the Population Association of America subgroup asked if DADS was the source of all the Bureau's products. Mr. Kavaliunas responded that some of the products would be available through DADS, while other data would be available through other mechanisms. For example, an electronic version of a Public Law file would be available through DADS or a user could order a paper copy of the same file.

Dr. Voss asked what downloadable files were. Mr. Kavaliunas explained that data users will be able to view a table via DADS via the Internet for free. If a data user wants to download multiple tables or multiple geographies, there would be a charge.

Dr. Voss believed that some of the prices seemed to be rather expensive. For example, downloads of quick tabs, priced at $312, would quickly become costly if the data user needed several data sets. This price would prohibit data users from using the quick tab feature, since it would be more cost-effective to buy the CD-ROM.

Mr. Kavaliunas said that the quick-tabs feature was priced for its convenience. Data users would be paying to have the data readily accessible through a download.

Dr. Stolzenberg of the Population Association of America subgroup said that the prices for data products would prohibit their sale. Dr. Voss concurred that product cost would prohibit individual use and encourage private companies to buy the census data and redistribute these data to customers free of charge or at a greatly reduced price. Following a quick calculation, Dr. Meyers of the Population Association of America subgroup added that it would cost him $8,000 to purchase the Public Use Microdata Sample (PUMS) files for Los Angeles County alone.

Ms. Becker of the Population Association of America subgroup asked why the Bureau was charging for its products when anyone could get the same information from the State data centers or a number of other commercial firms for free.

Dr. Voss asked why an "off-the-shelf" Public Law data CD-ROM was $150 and a custom CD-ROM was $30. Mr. Kavaliunas said that the custom CD-ROM price reflects the average sale cost. The majority of custom CD-ROM buyers will buy one CD-ROM with specific data. The average cost to copy these data onto a CD-ROM is approximately $30, which includes a $25 handling fee and 5 cents per area copied.

Dr. Spiro of the American Marketing Association subgroup clarified that data users buying the "off-the-shelf" product would be buying all the data. The custom CD-ROMS only include those data the data user requests.

Dr. Voss said he could not understand why it was less expensive to produce a custom product. In theory, a mass-produced product should be less expensive than a custom product. He believed the map manipulation and Geographic Information System (GIS) capabilities was an interesting idea; however, he expected that many private companies would offer the same service. The Bureau should be conscious of this competition when it prices these products.

Regarding online payment, Dr. Voss worried that when data users pressed a "yes" button they would unknowingly be charged for a product. Mr. Kavaliunas said that data users would receive a "pop-up box" with the specific cost of the tabulation or product identified, before they accrued a charge.

Dr. Spiro commended the Bureau for looking at its pricing in a systematic way. She added that pricing is one of the most difficult tasks in marketing a product. She believed that the Bureau must listen to its customers to determine the price of products. To determine which products the Bureau should make available, it will be important to know what products customers are using. If the products are available free from the State data centers, this should be considered when pricing the products.

Dr. Klerman of the Population Association of America subgroup said the real issue was not how much customers are willing to pay. The Bureau should be concerned with what customers are paying for the same data from alternative sources.

Mr. Kavaliunas stated that the Bureau is bound by Congress to provide the data at a minimum cost, but also must recover the cost of producing its data products.

In response to Dr. Spiro's question regarding the cost of producing the census products, Mr. Kavaliunas said the $52 million cost was for the 40 percent of products actually being sold over a 5-year period. The $52 million was the cost that could reasonably be passed on to the consumer. If the entire cost of dissemination were to be passed on to the consumer, the cost of Bureau products would quadruple.

Dr. Spiro said that from what she heard from the other subgroup members, the Bureau is driven by competition, since the information is available from other sources. Dr. Stolzenberg added that the census data were available cheaper, better, and faster from other sources.

Dr. Spiro said that there are other ways of pricing the products. The Bureau can price them to recover part or all the cost of production. Since the Bureau is competing with other companies who are putting out the same data products, the Bureau should be selling its products cheaper than the competition. If the competition is giving away the products, the Bureau should be doing the same. If the Census produces products that no one else offers, there should be an increased charge for these products. Those products that are exclusive to the Census Bureau should be priced higher to recover part of the dissemination cost.
Mr. Kavaliunas said the Census Bureau does offer products that no other source is able to provide—custom tabulations.

Following a discussion of a price list compiled by Dr. Spiro, Mr. Kavaliunas said that the pricing for CD-ROMS has not changed for the past 10 to 15 years. Ms. Becker said that the Bureau's prices have always been higher than any other Federal agency. Dr. Klerman said that the Bureau's high prices have encouraged other companies to buy Census Bureau's data and resell them for much less on the open market.

Dr. Stolzenberg said that because the Bureau is being undersold, customers are being lost. The Bureau can sell one CD-ROM for a very high price, or it can sell a lot of them for a low price. If the price is too high, no one will do business with the Bureau. An even more fundamental argument can be made—the Bureau depends on the support of the public to conduct its mission. If there are a lot of people using products the Bureau produced, a loyal customer base will be created. If the Bureau brings products to market that motivate data users to seek alternatives sources, the Bureau will lose its customers' loyalty.

Dr. Klerman said that if he were a small business man who had to fill out an economic census questionnaire and was then charged such a high price for economic data, he would be angry with the Census Bureau.

Mr. Kavaliunas said that he was confused—some members were telling him to raise his prices and others were telling him to lower them. Dr. Stolzenberg said the Bureau must raise its revenues by lowering its prices.

Mr. Kavaliunas said that in past focus groups, the Bureau was told that cost would not be a factor if the Bureau produced a product comparable to that being produced by the private sector. Ms. Becker responded that the Bureau's problem is that it does not know its own customer base. It does not know who is using the information at the State data centers.

Dr. Stolzenberg cautioned Mr. Kavaliunas that focus groups are a good way to get ideas, but a lousy way to conduct research.

Dr. Klerman warned that charging more than $20 for a CD-ROM will make it worthwhile for a private company to buy one of the Bureau's CD-ROMs and print them under their own name. Since the Bureau is unable to copyright its products, there are a lot of universities and public interest groups that would be happy to make names for themselves as data distributers. Currently, there are dozens of companies redistributing 1990 data. Many companies are giving this information away for free over the Internet, which is a very inexpensive way to make data available. This inability to compete with the private sector suggests that the Census Bureau should get out of the business of dissemination. All the offices that currently disseminate the data should be closed and the CD-ROMs should be given to two commercial firms (to preserve competitive pricing) for distribution. The Bureau should not be maintaining this $50 million endeavor when there are a number of companies that are happy to do this work themselves.

Mr. Kavaliunas said that Dr. Klerman's suggestion of outsourcing sounded like giving exclusive rights to a company for production of census products. The Census Bureau is not allowed to grant these rights. Ms. Becker disagreed with Mr. Kavaliunas. She noted that there had been a precedent for the Bureau granting exclusive rights to a company. Following the 1980 census, there was a special ZIP-code tabulation produced by the Bureau. Approximately 17 companies pooled there resources to purchase the exclusive rights to these tabulations for 18 months when the standard for exclusive rights at the time was 6 months. After 18 months,

the tabulations became a public domain product. [Note: The Office of the General Council at the Department of Commerce has since ruled that the Census Bureau should not grant exclusive rights for its tabulations.]

Ms. Becker stated that the Bureau's approach to providing data products was "strongly critized." The Bureau is not looking at how to get data into the hands of data users or what data are needed by the public. The Bureau must look at the needs of the data users and what products are required. She asked—Why does it cost $52 million to produce just 40 percent of the Bureau's products and what the Bureau was including in this figure? She wanted to know why the Bureau was charging such high prices when other agencies were selling their products for far less or giving the products to data users free of charge.

Mr. Kavaliunas responded that other agencies were not required by Congress to recover some of the money spent on dissemination. Dr. Stolzenberg responded that the Bureau will obviously not recover its cost very well considering its pricing.

Ms. Schneider summarized the points of the meeting as follows: the Bureau—(1) must broaden its base of support, (2) needs to guarantee broad access to data, and (3) must achieve the first two points while staying within the appropriations the Bureau will receive for data dissemination. She cautioned that, in the past, the Bureau has seldom received the money necessary to adequately disseminate data following a census.

Dr. Klerman said that these goals were fine, but they were missing the opportunity to save money. Internet access would be cheaper than printed products. The Bureau used to spend a lot of money printing and distributing paper copies of its data products. This should be stopped in favor of making the data available over the Internet. In addition, he seriously questioned whether the Bureau should be in the business of disseminating data. There are many private companies who are doing a fine job distributing the Bureau's data. If the Congress or the OMB require the Bureau to recover its production costs, then legislation should be passed that allows companies who primarily repackage census data to be charged a small fee. Following the Topologically Integrated Geographic Encoding and Referencing (TIGER) effort, an entire Internet industry was developed using the TIGER files to provide data users with map data, but the Bureau has yet to receive a penny from these companies. When there are companies directly receiving money from a census product, there is no reason why the Bureau should not recover some money from these companies.

In response to a question from Dr. Jacobsen of the Population Association of America subgroup, Mr. Kavaliunas said the Bureau can not receive royalties from companies using census data. Ms. Schneider added that a cooperative research and development agreement was an option. Theoretically, the Bureau could enter into a partnership with a private firm that takes a base product, adds value to it, and sells it to consumers.

Dr. Stolzenberg said that the Bureau does not get credit for its products. The Bureau should be able to put its emblem on every Internet site that helps people get data, including those sites that give people directions to their dentist's office.

Dr. Meyers asked if the Bureau could require the use of a trademark by those companies using census products. Ms. Becker suggested a trademark that would be recognized, similar to the "Intel inside" emblem.

Ms. Schneider said that certain Bureau products, including TIGER and the *Official Statistics*, are trademarked, but the Bureau can not require companies to include them on their products.

Dr. Voss said that he agreed that the Bureau should be given recognition for its data products when they are repackaged by other companies. His one reservation was that the Bureau will lose control of data quality. Any incorrect data may reflect poorly on the Census Bureau.

Dr. Stolzenberg lamented that his service with the subgroup has been a saddening experience. In his many years working with the Bureau, he has developed a real affection for the Bureau and the people with whom he has worked. However, the Bureau is terrible at promoting itself.

Dr. Etzel of the American Marketing Association of America subgroup said the Bureau knows its inability to compete with companies giving away census products. The subgroup should be talking to the Congress, not the Bureau, since Congress needs to pass legislation before the Bureau can do anything to make its products more competitive or at least receive credit for its data.

In response to Dr. Klerman's question regarding the $52 million dollars cost of data products, Mr. Kavaliunas said that the money was used to pay employees to put together tabulations and review the census tables. Dr. Klerman suggested the Bureau think about getting out of the business of dissemination and leave that business to private companies.

Dr. Voss said it was obvious the price list had not been changed for years. For example, the price charged for a CD-ROM has remained the same for years. The cost of producing and copying CD-ROMs is much less expensive today than it was 10 years ago. Mr. Kavaliunas said that the cost of CD-ROMs reflected the cost of the staff needed to produce that software.

Dr. Klerman reiterated that the Bureau needs to reengineer its budget, which would include the removal of a large numbers of employees. The Bureau should create tables once, put them on the web, and let the public have access to them. The subgroups' job is not to protect the Bureau staff; it is to protect the collection of data. The $52 million should go into data collection. Since so many members of the Census Advisory Committee of Professional Associations are complaining that the Federal statistical program is "going to hell in a handbasket," it should be obvious that this $52 million is not being well spent. Given the fact that technology and private companies are able to deliver these data in ways that customers are willing to pay a lot of money for, it is obvious that the Bureau is not allocating its tasks well.

In defense of tradition, Dr. Meyers said that when he sees students get a census CD-ROM from the library, they are looking at a very shallow slice of the data when compared to what they could get from the published census tract books. Dr. Klerman responded that published data would not be needed if a better DADS system were constructed to provide the information the published census tract books provide.

Dr. Meyers said the quality of the data is in its use, not just the validity of the raw material. He is concerned that people can not browse the volumes of material available when

they are online the way they can with the printed data. Data users can perform a search much better online, but they can't see the row headings, column headings, or the context. As a result, there is a degradation of value.

Ms. Becker said that she is concerned that less skilled people are looking at "unfinished" data—the raw demographic data provided in the Summary Tape Files. Lay people often do not know where the data they are looking at came from or what it represents. Dr. Stolzenberg asked Ms. Becker if she wanted people to receive a prescription from a licensed demographer prior to looking at these data. Ms. Becker responded that there is an obligation among licensed demographers to create finished products for people.

Dr. Klerman said that if the Congress wants the Bureau to provide a finished product, then the Bureau should. In the absence of that decision, the Bureau should not be producing finished data, and it certainly should not be trying to recover its cost in the way it is because the Bureau will not succeed. The agency must tell the Congress that these products are valuable, and they must appropriate money for them. The Bureau has to decrease its cost, or it can not produce the products.

Dr. Jacobsen said that from her own experience, the cost of data around the world is very high. The subgroups' mentality that all this information should be available on a $20 CD-ROM does not exist anywhere else. It may be useful for the Bureau to use the experiences of other agencies to illustrate to the Congress that the legal requirements the Bureau must follow are crazy.

Dr. Voss said data users are so use to getting census data for free that no one is going to put a price tag on them. Dr. Stolzenberg said that he already pays for the data through his taxes. People are already "homicidal" after completing their census questionnaires. It isn't right that the Bureau require survey recipients to provide their information free of charge, but the Bureau requires payment to purchase the final data.

Dr. Spiro said that if the information is on the market for free, then the Bureau can not expect to compete. The Bureau needs to be able to copyright its data products if it wants to make money from their sale.

Mr. Adams of the American Marketing Association subgroup said that the real issue is one of pricing. Whether or not the Bureau will continue to be in the business of data dissemination is a political issue. He understood if the Bureau employees present were frustrated that the subgroups had not given any substantiative advice aside from the advice that maybe a logo, like the "Intel inside," be placed on all census products.

Dr. Stolzenberg reiterated that the two subgroups were telling the Bureau that the pricing of census products was much too expensive. Mr. Kavaliunas again said that he was hearing two opposing opinions—the prices were too high and they were too low.

Dr. Klerman said that the prices were obviously too high. The number of projected sales the Bureau provided for its products were trivial considering the millions of hits private Internet sites providing census data are getting every day. The Bureau is talking about thousands of hits in 5 years. If the products are priced high enough, the number of orders can be counted in the thousands; however, there is business for these products in the millions.

Ms. Becker said that the Bureau doesn't understand the dissemination of its own products once they are made public.

Dr. Voss asked if it would make sense for the Bureau to hire a contractor to design a marketing plan or Census Bureau products.

General Edit and Imputation Research (ASA)

Dr. Winkler said that the Bureau is trying to develop generalized processing systems because they can foster greater efficiency through being used on many surveys—particularly edit and imputation systems—and because institutional knowledge can be incorporated into the software. The two systems under development are the Structured Programs for Economic Editing and Referrals (SPEER) and the DISCRETE systems. The key underlying facet of these systems is the Fellegi-Holt model of editing. The edits reside in easily modified tables that enable users to translate the edit rules into tables that are easily maintained. The source code is reusable and needs no modifications and checks the logical consistency of the system prior to production editing. In the past, this was very difficult to check, but is automatic in the Fellegi-Holt model. The model shows that if a user keeps track of "implicit edits" the records being edited can be corrected in one pass through the system.

The key to the Fellegi-Holt model is the generation of implicit edits, which can be derived from the explicitly defined edits. In the past, records being edited could pass explicit edits in the original pass, but then failed when something had been changed and the record was reedited. In the Fellegi-Holt model, the implicit edits can assure that the changes made to the record will still result in satisfying all edits.

The Fellegi-Hold model defines an edit as a set of points; it may be a set of points that show a record for a person who is married and is less than, or equal to, 15 years of age. In the U.S. that would normally be considered an unacceptable relationship and the record would fail an edit. If explicit edits imply other edits, then when the implicit edit fails, at least one of the explicit edits also would have failed. The implicit edit provides the information that assures that the record can be changed to something that will satisfy all edits, and within that capability, allows the minimal change needed. To move from the explicit edits to the implicit edit, an editor would generate on a specific field, then take intersections of the remaining fields.

The generalized economic editing system has been successfully implemented in various Census Bureau surveys. This system uses ratio edits and used a fairly sophisticated hierarchy that enables it to be adapted rapidly to different surveys. Typically, with economic files, users want things to "add up." There are edit systems that attempt to deal with this requirement, but no system can assure that records will add up. Last year Bureau staff developed some procedures to satisfy both the ratio edits and balanced equations.

For the DISCRETE edit system, users face an integer programming problem. The real killer has been the generation of the implicit edits; it is a full scale set-covering problem of the hardest type. To deal with it, the Bureau introduced a system in 1995, but new algorithms have been developed since that enable the system to "learn" which computational paths were successful, then eliminate computational paths that were less successful. This resulted in a very substantial improvement in the speed of edit generation (from 6 hours to about 2 minutes for the file corresponding to the decennial short-form questionnaire). In consultation with Italian

statisticians, the Bureau ran new algorithms on all of the subcomponents of the file. The program ran in 28 hours and successfully generated 99 percent of the edits[2]. (The reason the system cannot generate all of the implicit edits is that when a survey has skip patterns the system cannot "learn" at the beginning stages of the computation as efficiently as hoped.) The key point is that the increased edit speed will enable the use of certain new types of production systems. For example, the agency is looking at developing a census short-form system. Currently, the Bureau has an essentially research system, which takes the same inputs as the decennial system and produces outputs in the same form. The research system is, in fact, a full production system and could be run parallel to the decennial system. By using sophisticated conversion techniques, the software has been made to run much faster in comparison to previous systems of this type. Until recently, no one could implement a Fellegi-Holt system on a census file for such things as an age edit. (The Canadians tried; they found they had to use decade instead of year as the measured unit.)

On the imputation side, the Bureau has a great deal of experience and has developed very explicit statistical models. These use log-linear models for the housing characteristics and Yves Thibaudeau at the Bureau developed and introduced a number of techniques that have been very successful. For person characteristics, Tod Williamson developed logistic regressions for such items as relationships to head of household, and ordinary regressions for age characteristics and differences in ages. The system adapts automatically at the tract level, uses all the characteristics available, deals with multiple-variable characteristics, and enables the Bureau to estimate imputation variances directly. None of this would mean much if it was purely experimental. The Bureau, however, has developed awesomely fast algorithms—on the order of 100 times as fast as those used in the past—which means the agency has a potential production system that runs as fast as the existing "hot deck" system. The agency can now process a district office file in about 2 hours. The systems planned for use in Census 2000 will be about 10 times as fast.

A critical factor is that the comparisons studies done with data from different parts of the country show the new system outperforming the previous hot deck system. In areas with heavy concentrations of minority populations, the estimates are consistently better than those produced using the older system.

The Bureau has three major questions for the members to address—

- Should the Bureau continue to develop generalized systems or should the agency try to purchase generalized software? If the latter, why?

- Do members have any specific suggestions about the direction edit/imputation research should take? Are there any additional areas the Bureau should consider?

- Do the members have any suggestions regarding developing additional edit modules, better user interfaces, and creating suitable documentation and training methods?

---

[2]The Italians had to break the files into four subcomponents. When they attempted to edit all subcomponents simultaneously, their programs did not complete in 8 days on the larger IBM mainframe.

Dr. Stasny said the members had received three papers as background for this presentation, one each on the SPEER and DISCRETE general editing systems, and the third on modeling for imputation on the Census 2000 short-form questionnaire. The SPEER system is a program developed for editing continuous data subject to balancing equations and ratio constraints. The technique suggested to increase the speed of the algorithm was to use simpler and faster codes that might partially correct a record in one pass and use several passes to complete the edit. This seems a good idea, and the results shown in the paper indicate that the SPEER system produces edited records that satisfy the constraints of the program in a few passes, with significant improvements in speed. What was not shown in the paper was the quality of the data produced.

The DISCRETE system is used for editing data from discrete surveys with no skip patterns, and the paper gave mostly theoretical results that the algorithm worked. The results showed huge reductions in central processing unit time using the DISCRETE system compared to other editing systems, but again, the question of the quality of the edited data is not explicitly addressed. DISCRETE is a work in progress, but looks very promising.

The paper on item imputation indicates the Bureau is taking a modeling approach that uses selected housing-unit items and variables like those used in the hot deck processing. A log linear model is proposed for each tract with three-way interactions among all the housing unit items, two-way interactions with race and the race of the preceding responding housing unit, the tenure and the race of the preceding housing unit, and the three way interaction between tenure of the current, previous, and following housing units. These are hierarchical models that include all the lower order interaction terms. The algorithm used seems to use a missing-at-random (MAR) missing data mechanism. The Bureau then used the multinomial probabilities corresponding to the cells of the table to generate imputations randomly according to those probabilities.

“Person items” on the short-form questionnaire that have to be edited and imputed could be relationship to the householder and age of the individual. A series of logistic regression models are used for the relationship to the householder item while a standard regression model is used to impute age. The paper indicates that the systems have been tested on 1990 census data from various locations, and improved race imputation and slightly the average age on imputed values. She encouraged continuing research on this approach.

There were several issues the Bureau should address. Evidently the models will be used at the tract levels and different parameters in the model will be estimated for each tract, while the form of the model would remain the same. The Bureau might consider allowing for using different models for different parts of the country where various data items will be considered of differing importance. The Bureau would have to fit a lot of models at the tract level, but the new algorithms are running so fast that fitting in that many models might be a problem. Variance estimation was not mentioned in the papers, but estimates that take missing data into account can be obtained almost immediately.

Dr. Stasny said it is unclear whether the new imputation systems will be used in Census 2000, and she would like this issue clarified.

The Bureau should continue development work on these systems; the possible benefits from cost savings and in higher quality data are substantial. She assumed that if the required

generalized software systems had been available the Bureau already would have purchased them, rather than trying to develop it inhouse. If such software was available, it would be the way to go, but since it evidently is not available she saw no purpose in urging the agency to buy them.

The modeling approach for imputation seems useful and the Bureau should continue to work on this methodology. She noted that her organization had used a Gibbs sampler on a sample from the economic census data and found it to be very time consuming. She liked the sampler, and urged continuing work on the system, but was not sure it is ready to be used for production.

With regard to the last question asked, the Bureau obviously needs to have a good user interface, as well as the best documentation possible.

Dr. Stasny congratulated the Bureau on the work done on the edit and imputation systems and said they looked very promising.

Dr. Winkler emphasized that the generalized edit and imputation systems he has described are still research projects; the hot deck system will be used for Census 2000. The generalized systems may be run parallel to the production system in the Census 2000 Dress Rehearsal. He noted that the results of using the generalized systems have also helped confirm that the hot deck system works very well overall for univariate statistics.

With regard to the SPEER system, Dr. Winkler reported that the Bureau compared the SPEER system to the Statistics Canada's Generalized Edit and Imputation System (GEIS) several years ago using simulated Canadian agricultural data for the test. His own assessment was that the GEIS was the better system, although the SPEER system performed better with that particular data. In each case, the default modes in the systems produced very high-quality data. He pointed out that what drives edit imputation is the imputation, rather than the editing. If a good job is done on the imputation, then the output microdata will be very good.

Turning to the DISCRETE system, he said this is the first work the Bureau has done in this area. It has become clear over the last 10 years that the imputation methods used by the statistical agencies are out of date and inefficient compared to other methodologies available. Accordingly, the Bureau decided to try some of the newer systems and discovered that the new systems make a significant difference, particularly when dealing with multi-variable characteristics. The work the agency has done with the new system reveals that enormous amounts of data can be handled very efficiently and with high quality results.

In reply to a question by Dr. Tourangeau, Dr. Winkler said that, on average, about 22 percent of household records and 2 to 5 percent of individual items would require editing. Most of this will actually be imputation, since the overwhelming majority of editing problems will involve missing data. He noted that the sheer size of the data structure that is involved presents an enormous challenge and required some very clever conversions to enable the systems to work.

Dr. Bell said he found the imputation paper particularly interesting because he has spent a fair amount of time thinking about imputation for a project on which he was working. He became concerned about the assumptions in the various models, which led him toward what

might be called model-based hot decking. This involved some modeling used to define cells from which to draw a random case. That made him a little more comfortable about not having to check model assumptions. He said he also tried to simplify things by trying to impute only one variable at a time, and using that variable as part of the input for the next imputation.

Dr. Winkler said that the Bureau had looked at a very large number of closely related models, and tried to identify those that gave the best predictive capability for imputation. Looking at the way the hot deck model was designed provided useful insights into how to go about obtaining good "donors" or good prediction in imputation. The agency has done extensive modeling and exhaustive testing of those models. Cross-checking on new procedures is critical, which is why the Bureau plans to run a generalized system in parallel to the production system used for the dress rehearsal; both systems can be run simultaneously and do detailed comparisons very rapidly. The Bureau's systems have been designed to enable different scenarios to be checked—a test can be plugged in and they can produce numbers to be checked in less than an hour. There has been a great deal of good work on these systems done in the past and the agency is anxious to build on that work.

Dr. Bell commented that the Bureau's use of the missing-at-random assumption in imputation for relationship has led to imputing relationships other than "spouse" or "child." This seems to make sense as it is more likely that a more complicated relationship will be left blank than a more conventional one. Dr. Winkler said the Bureau has observed a lot of variation in reporting relationship among district office areas, tracts, even among blocks, and thus has had to establish systems that will take those variations into account. The generalized systems seem better able to adapt to these situations than the hot deck system.

In response to a question by Dr. Tourangeau, Dr. Winkler said the Bureau does not carry out any sort of retrieval interview program to check the accuracy of its inputed data. The agency can obtain two sets of estimates—one from the hot deck system and a second from one of the newer generalized systems. Often, particularly in areas with high concentrations of minority populations, the estimates are far apart. In determining the reasons for the differences the Bureau has found that a frequent problem is the fact that the hot deck system does not have the right type of donors. This finding is confirmed by similar results in Canada and the United Kingdom. Experience shows that agencies have to be very careful how they model using the hot deck. The British have been able to dramatically improve their data by improving their modeling. A critical factor is the need for a good pool of donors; it is very difficult to assemble a really good one.

Responding to a question by Dr. Wilson of the Population Association of America subgroup, Dr. Winkler said that the Bureau has not done very careful investigation of error generated in relationships generated by multiple imputations. In general, the variation for most household characteristics is very small. The Bureau has not analyzed items such as age and relationship, but it is obvious that the age estimates are consistently coming out younger than those produced using the hot deck. The Bureau's Population Division staff and others have known for some time that the estimates being produced for age were too high, but they did not know how to approach the problem. When successive modeling is done, the modeling becomes a better and better fit to the actual data, which is promising for the improved accuracy of imputed data.

Chief Economist Updates (AEA)

Dr. Haltiwanger updated the subgroup on the recent activities of the Center for Economic Studies (CES). He noted that the planned expansion of the research data center (RDC) program is continuing. In December 1997, the Bureau issued a call for proposals for new centers, using every conceivable means for distributing the information. The National Science Foundation (NSF) is participating in the activity, not only by helping to transmit the call to individuals or institutions likely to be interested, but also by providing peer review resources for the Bureau's consideration of the proposals. That review of proposals is underway now and will be going on for some time. The Bureau expects to make its initial decisions regarding the sponsors for the new centers late this spring. The agency plans to inaugurate two new RDC's this calendar year, and two more in 1999. The next round of proposals will be due on August 15, 1998.

In response to a question by Dr. Willis, Dr. Haltiwanger said the Bureau will continue to consider proposals submitted for this year's new RDC's until it finds two that are satisfactory.

Replying to a question by Dr. Scherer, Dr. Haltiwanger said the NSF has committed up to $500,000 per year as "seed money" for the new centers—for any given center $100,000 per year in supporting funds for the first 3 years of operation.

In reply to a question by Dr. Gort, Dr. Haltiwanger said the Bureau has not decided on a total number of RDC's. A significant problem in expanding the program is the agency's ability to support additional centers. The three centers currently operating require a good deal in terms of agency resources, including Bureau staff on site. Managing access to the Census Bureau's data from the RDC's also is a cause for concern.

Responding to questions by Dr. Scherer, Dr. Haltiwanger said each RDC must have a secure site—i.e., a secure room with controlled physical access, a considerable amount of equipment, no electronic access to the outside world, and 24-hour access for researchers—with a full-time Census Bureau employee onsite to handle interaction with researchers. The basic resources needed are the physical space, the hardware (computers, etc.), and the onsite Census Bureau employee (the Bureau employee is not there to act as a security guard but as a liaison with researchers). Also, there are other expenses involving what may have to be done to support a given research project by Bureau headquarters staff, such as special data extracts.

When a researcher wishes to use an RDC, the candidate must first submit a proposal for local and Census Bureau headquarters review and must have funding to support the project. If the proposal is approved, the Bureau may have to provide some data extract—frequently done at the Washington, DC, headquarters—and provide that extract to the researcher. Since there is no outside electronic access to the RDC, the data extracts or other electronic data files from Bureau headquarters have to be copied onto computer tape or discs and transported physically to the RDC site. The onsite Census Bureau employee's responsibilities include familiarizing researchers with the agency's data sets and helping them use the data to meet their research needs. For projects that may involve microdata that have not been developed for internal use, such as Survey of Income and Program Participation (SIPP) data, there may be some difficulty in providing access until further data development work is done. For projects that require special access or service and support activities are needed, the Bureau is considering requiring a user fee.

Replying to a question by Dr. Lilliard, Dr. Haltiwanger said many people have expressed interest in the RDC program and the Bureau has received more proposals than can be accommodated by the planned expansion of the program. He noted that the quality of the submissions is very high.

Turning to other activities at the CES, Dr. Haltiwanger said the Bureau and the CES are interested in determining what new kinds of databases can be created, and what sorts of access can be given to users. The CES is trying to link establishment economic data with demographic data files. The Bureau is in a unique position to do this since it has the universe files on both businesses and households and can bring the data together in detailed geographic ways, or even link employer and employee data sets. The possibilities for research are enormous—a matched longitudinal employer/employee data set is the "holy grail" of labor economics.

Responding to a question by Dr. Scherer, Dr. Haltiwanger said the Bureau can match an individual employee to a specific firm. The agency already is working on studies of healthcare issues and the impact of other benefits on worker quality and effort. There is considerable enthusiasm at the Census Bureau for this sort of work. The agency is jointly sponsoring—with the Bureau of Labor Statistics, the Sloan Foundation, the World Bank, and others—a conference on May 20-21 on linking employer/employee data. The sponsors have been struck by the range of individuals and organizations both in the United States and abroad that have expressed interest in participating. The Bureau was particularly interested in having foreign participation, since some European and other statistical organizations and agencies have done a great deal of work in this area.

Confidentiality of data remains a major consideration in any plan for matching these data sets. Some of the files being used contain individuals' address information, and the creation of some combined data sets involve the use of highly sensitive administrative information.

In reply to a question by Dr. Scherer, Dr. Haltiwanger said the matching that has been done has used administrative records data and the Bureau's demographic and economic data files.

Dr. Haltiwanger pointed out that the Census Bureau is involved in developing the protocols for creating and using these data sets. All of the data are inside the Bureau's "firewall" protecting the confidentiality of the records. The agency also is trying to keep all the files involved in one location and on a limited number of computer terminals, and is very cautious about access to the files. There is a great deal of enthusiasm about working with these files, but it will take some time to work out the procedures needed to protect the data while using the files for the maximum benefit. The Bureau welcomes members' ideas for working with the data, as well as about controlling access.

With regard to access, Dr. Haltiwanger said that in the RDC program, individual researchers must agree to uphold the confidentiality of the census data (violators are subject to criminal penalties) and there is limited physical access to the data files. The Bureau is interested in whether the penalties established for violating confidentiality are sufficiently severe, and whether there are other things the agency can do to promote the confidentiality of the data.

Dr. Scherer suggested the Bureau consider doing background checks on researchers applying to the RDC program to try to identify and conflicts of interest. Dr. Knickerbocker pointed out that the Bureau does a general background check on applicants.

Dr. Willis commented that most discussions of confidentiality center on technical issues rather than on the motivation of someone to violate the requirements. The Bureau has to guard against that happening. The RDCs are set up well, in some ways, to protect against those sorts of violations, but he was not sure that the "technical firewall" the Bureau had erected to protect the individual demographic and economic data is as efficient.

In reply to a question by Dr. Willis, Dr. Haltiwanger said the Bureau has not yet had any researcher participating in the RDC program involved in any violations of the agency's confidentiality or other requirements.

Dr. Willis said he also has done work using Social Security Administration records; the arrangements his organization made required that researchers have a grant from some Federal agency. The implicit sanction was that anyone who violated confidentiality of the records would never get another grant. He noted that he has never heard of a breach of confidentiality within the academic research business. This made discussing the problem seem a little strange since he had no real experience with the actual situation.

Dr. Lilliard commented that the Census Bureau needs to think through its rules before imposing additional restrictions. Sooner or later there will be researchers who will have some sort of links with firms whose data are in the files the researchers are using.

Dr. Scherer pointed out that at the Federal Trade Commission (FTC), the general rule applied was that if a staff member had access to the line of business data, he or she could not take part in any FTC litigation proceeding underway at that time. He added that when he was a Census Fellow he was an independent consultant in a litigation involving Northrop, and he asked the Bureau at that time to strip off any Northrop data from any files to which he had any access.

Dr. Willis noted that a screening process will necessarily eliminate some researchers. The question the Bureau needs to address is whether it can make agreements with researchers that will work appropriately. He wondered if it is possible for a researcher to obtain data through an RDC for second party, and how the security rules would apply in such a case.

Dr. Haltiwanger said that graduate students are, in fact, most often the researchers that are working at the RDCs.

Dr. Maynard suggested that some of the problems could be eliminated if the matches were run by someone whose loyalty was to the RDC or to the Census Bureau, rather than the researcher.

Dr. Haltiwanger noted that the employer/employee data matching mentioned earlier will not be done by individual researchers in any case. All that is done at Bureau headquarters.

Dr. Knickerbocker commented that his concern is not the Bureau's exposure to malicious intent, but the possibility of academic enthusiasm leading to the "innocent" or unintentional release of confidential information.

Dr. Haltiwanger commented that table 2 in the annual report for the CES shows the kinds of data sets being produced at the Census Bureau. On the table, a system of star annotations is used, with one star indicating sets with which the Bureau has some experience using for matching, and which are available now at the RDCs, and two stars indicating "core" data sets available at no additional cost to users. Those sets with no stars indicate data sets that have not been used in this fashion.

He noted that many users are interested in using data files from the population censuses. The Bureau is very interested in obtaining sponsorships for big data sets suitable to analytical purposes. The agency has a program of this kind currently underway—sponsored by the Immigration and Naturalization Service—to develop a microdata set from the 1960 through the 1990 population census data files. This project involves putting together a SAS data set for the full 100-percent data file, plus the 1-in-6 sample file, and will take considerable staff time, with a projected completion data sometime in 1999, but once completed will be available to other researchers as well .

Dr. Haltiwanger pointed out that one question the Bureau needs to address is how to put together the additional resources to do the kinds of data set development needed for the RDC program and other matching projects.

In reply to questions by Drs. Willis and Lilliard, Dr. Haltiwanger said that the demographic data files have specific problems, having to do primarily with historical consistency of the files. Some of the internal files, such as the SIPP files, are difficult to use because the public-use files released are in a different format than the internal workings of the files. The internal versions have inconsistent formats and no documentation.

In response to a question by Dr. Betancourt, Dr. Haltiwanger said the Bureau plans to deliver two major data files to the RDCs this year—an economic census file for the 1982, 1987, and 1992 censuses with longitudinal linkages, together with the Bureau's current Standard Statistical Establishment List (SSEL) file. The Bureau is also working on a longitudinal SSEL file covering the years 1987 through 1996.

Dr. Maynard wondered why there seems to be an either/or situation with regard to data files—researchers can either use the public-use microdata files or go to one of the RDCs to use the raw data files. It seemed there is a need for some sort of intermediate data set with the critical fields needed that did not make it to the public-use data set because of confidentiality considerations. She suggested that such an intermediate file, together with the public-use file, would fulfill most researchers' needs.

Dr. Haltiwanger said something of that nature could be done for the Current Population Survey file, and while there will be problems with the decennial population census data, it might also be possible to do something there as well. However, some files are less likely to be successfully adapted in this fashion.

In reply to a question from Dr. Lilliard, Dr. Haltiwanger said there are two ways for a researcher to go about obtaining more usable data from the raw SIPP file. First, the researcher or organization could pay for a research assistant to come to Bureau headquarters to help assemble the part of the file needed. Alternatively, the Bureau and the researcher could put together a project, and find the necessary sponsor or sponsors, to work over the entire SIPP file and create the documentation in a way that the file could then be made available to the RDCs.

Responding to questions by Dr. Willis, Dr. Haltiwanger said it would be a good idea for the Census Bureau to put together public-use data files from the economic data files, but little has been done in that regard so far.

Dr. Maynard encouraged the Bureau to impose a condition for the use of the RDCs that when data sets are created that meet the standards for public use, those sets are retained by the RDCs for the use of other researchers. Dr. Haltiwanger said that this is already the practice at the RDCs, and that the Bureau asks researchers to prepare documentation for their data sets as well, although the quality of the documentation prepared is uneven.

Dr. Willis added that SSA records include a great deal of data on individuals' work history with different companies, but SSA will not release this information on confidentiality grounds, although they are willing to permit researchers to use a data set that includes information on the time workers are at a company, providing the name of the company is not revealed. Would the RDC entertain projects the purpose of which was to produce a public-use data set from an RDC data set that would otherwise remain confidential?

Dr. Haltiwanger commented that this is an interesting idea. It is possible that new kinds of public-use files can be created by linking existing data sets, providing the confidential data in the matched sets were protected. There did not appear to be an inherent confidentiality problem with that sort of project.

Dr. Willis suggested there are two possible ways such work could be done—the Census Bureau itself could produce public-use data sets and could undertake such a project, or something similar might be sponsored by the National Science Foundation or some other agency.

How Should We Promote Confidentiality in the Decennial Census? (AMA)

Census staff made presentations on the Bureau's three-part approach on confidentiality messages—Ms. Waldrop's presentation was against using advertising, Mr. Gates' presentation was in favor of using advertising along with partnerships, and Ms. Marks' presentation was on what exactly the Bureau has been doing for the Census 2000 Dress Rehearsal.

Ms. Waldrop emphasized that the agency considers the confidentiality issue very important. The Bureau is particularly concerned because of declining response rates in not just the decennial census but also in all types of surveys. The agency has done much research on what the confidentiality message should be, but no research has been done on the impact of the messenger. She noted that this presentation was about who should be delivering the confidentiality messages and presented the case for promoting confidentiality through partnership, not advertising. She believed that people who do not trust the Census Bureau will not believe anything in its paid advertisements. It would be preferable if the agency used

outside spokespersons (such as local “gatekeepers”) for spreading the confidentiality message. Including the confidentiality message in general advertising may raise concerns among audiences who are not currently focused on this issue.

Ms. Waldrop noted that the Bureau experienced during its 19 or so focus group meetings that people were easily agitated by the confidentiality message. The Bureau needs to reach every resident in the United States; therefore, can an advertising agency be responsive to all these different viewpoints, especially when conditions can change rapidly in small local areas? The agency has many partners representing many different populations, various race and ethnic groups, community leaders, and special interest groups who know the benefits of census and can more effectively deliver messages on confidentiality.

Mr. Gates was convinced that the confidentiality issue needed more than the partnership efforts. There needs to be a centralized campaign to deal with the privacy message. He did not believe that the public would react negatively to an advertised message that stressed confidentiality and noted that the environment for Census 2000 was quite different than that in the 1990 census. Networked computers, as exemplified by the Internet, have changed everybody’s focus towards protecting their respective personal information. He also noted that even businesses have faced up to the fact that privacy is an important issue that needed to be dealt with. Businesses are now promoting privacy and protection of information. Mr. Gates believed that the Bureau can learn from what businesses are doing in terms of advertising and incorporate some of that into its advertising campaign. He noted that if the agency relied solely on partners for delivering the confidentiality message, he would be concerned about the coverage and consistency of the message. A national advertising campaign can focus on the confidentiality issue successfully. Partnerships should not be the only approach.

Ms. Marks showed a video on television, radio, and print advertisement campaigns that are being used in the Census 2000 Dress Rehearsal sites. She noted that these advertisements were developed within a very short time period by Young & Rubicam and there were a few confidentiality messages in these advertisements. Based on research, Young & Rubicam decided that confidentiality messages would be for those groups for whom concerns about confidentiality represented a barrier to participation in a census, e.g., in the Spanish-language advertisements. Ms. Marks noted that confidentiality was a major concern among the Spanish-speaking population. She gave a handout that was prepared by Scholastic as a part of the Bureau’s “Census in the Schools” project under the advertising contract for the dress rehearsal. This take-home brochure includes a short privacy message on the second page. She noted that this brochure did not come across as an advertisement, but it had the appearance of coming through a trusted agent like a school teacher. Therefore, it was like a combination of partnership and advertisement. During the prewave of the Advertising, Marketing, and Partnerships Efforts Evaluation Survey, there were questions about trying to measure knowledge and concerns about confidentiality issue in the census. The same questions will be asked again during the postwave of the survey. She informed the American Marketing Association subgroup that Young & Rubicam would be doing more research on this issue to identify what would be the most salient way to address this issue and if there is a concern about this issue.

Mr. Adams agreed that confidentiality was a very important issue and commended the Bureau for its efforts to deal with it. However, he noted that he felt somewhat disconnected reading the Bureau’s paper and hearing the presentations, especially the mention of “declining

response rates to censuses and surveys may be linked to increasing concerns about individual privacy" at the beginning of the paper and "privacy message is a small part of the advertising plan... based on Young & Rubicam focus groups. The benefits messages are the most compelling while the confidentiality concerns are not generally an issue" near the end of the paper. He questioned why the confidentiality issue is a small part, and he did not think that the Bureau could send the advertising people off with one message and send the local partners off with another message. He believed the Bureau needed to have synergy between all communication channels. He suggested the Bureau quantify the role of confidentiality in the nonresponse rate. As a marketer, Mr. Adams said he would like to know if 80 percent of the problem in nonresponse can be directly attributed to the confidentiality issue. There has been too much reliance on qualitative research on this issue at the Bureau. He noted that the use of focus groups is still developing, and, therefore, using focus groups to make conclusions would be extremely dangerous for marketing research. Focus groups are not a reliable and valid way to measure issues. However, he applauded the Bureau's mandate to simplify the message.

Mr. Adams believed that the benefits message "what's in it for me" should be the primary message; however, there needs to be a more broad reassurance message on confidentiality than using the message only in Spanish-language messages. To reiterate the "disconnection," Mr. Adams said, if the goal for advertising is to increase the response rate, and the perception of confidentiality is correlated to the response rate, then advertising has to address confidentiality. He added that focus groups are of value, however, he did not think they could reliably and validly answer the questions being raised.

Mr. Adams discussed the three questions (see background materials) presented by the Bureau on the confidentiality issue. He noted that the American Marketing Association (AMA) subgroup members felt that the lack of a control cell in the dress rehearsal would prevent researchers from isolating the effects of advertising. On behalf of the members, he strongly recommended that the Bureau ask Young & Rubicam to undertake pretesting of the advertising campaign not with more focus groups but with other tools. Pretesting can provide answers to the Bureau's questions; focus groups cannot.

On whether partnerships are the best vehicle to convey the confidentiality message, Mr. Adams did not think this was a question of "either/or"; the Bureau should not send one army with only the benefits message and the other with the confidentiality message.

Ms. Marks mentioned some statistics from the Roper's poll asking "Why would you fill out the census form"? Sixty-four percent of the respondents said the benefits message was the reason and fifty-nine percent said because the information was confidential.

Mr. Adams still believed that too many focus groups were dangerous and useless.

Ms. Ashcraft agreed with Mr. Adams. She felt that 60 focus groups were too many. She noted that focus groups are used to develop and refine methods, but not to test. Ms. Ashcraft was dismayed to see the person on the video tape say that the commercials had been tested in focus groups. She too suggested pretesting of the commercials.

Dr. Etzel said, given the information the subgroup had, he was trying to understand why Young & Rubicam was approaching the confidentiality issue this way. He agreed that confidentiality was a problem; however, when he looked at the $100,000,000 budget, he did not

think it was enough to accomplish much. He believed, with this budget, that Young & Rubicam could only work with awareness and reminding people about census to increase the response rate, but not with confidentiality. According to Dr. Etzel, there are two important factors about dealing with confidentiality—(1) it would have to be highly targeted to the audiences that are concerned about confidentiality and those audiences would have to be identified and (2) it would have to be a very sustained campaign because the agency is trying to change their attitude. He did not believe that Young & Rubicam had the time frame or the dollar amount to do so.

Dr. Spiro added that she would like to see the confidentiality message in every advertisement even if it is only a blurb. She noted that she did not pick up the confidentiality message in the video tape shown earlier. She believed that most people had faith in the Government and, therefore, adding the phrase "by law" in the confidentiality message might make a difference.

Several members of the AEA subgroup discussed the effects of the use of words such as "privacy," "secret," etc. in the confidentiality message. Ms. Ashcraft emphasized that the message should be simple rather than complex.

Ms. Becker of the Population Association of America (PAA) subgroup noted that the Bureau definitely needed to test the advertisements to understand the impact on the audiences, and the agency should do some research to figure out where to focus.

Ms. Waldrop asked if the Bureau needed to mention the 72-year limit on the confidentiality of census information. The members did not feel the agency should mention that unless it was legally required. Mr. Gates mentioned that the Bureau included this information on the questionnaire in 1990. Dr. Etzel and Mr. Adams thought that putting qualifications on confidentiality in our outreach messages would only complicate things.

Mr. Cooper asked, when an advertising message mentions confidentiality, would that suggest the possibility of a confidentiality problem? Dr. Etzel said it was not possible to know who would or would not feel that way, and he felt uncomfortable to stereotype the Hispanic audience with the confidentiality message as was implied in the census video tape.

On behalf of the American Marketing Association (AMA) subgroup, Mr. Adams reiterated that the Bureau should ask Young and Rubicam to pretest the commercials.

How Will the OMB Proposal on Tabulation of Race and Ethnicity Data be Implemented in Dress Rehearsal Tabulations? (ASA/PAA)

Dr. del Pinal summarized the main changes in the classification of Federal data on race and ethnicity introduced by the publication of the Office of Management and Budget's (OMB's) *Federal Register* notice in late October 1997—

- Respondents were allowed to report more than one racial group.

- The Asian and Pacific Islander population group was divided into two, separate groups— Asians, and Native Hawaiians and other Pacific Islanders.

Several other recommendations for changes to the racial and ethnic categories were made but the OMB did not adopt them. The OMB also recommended that surveys and censuses use respondent self-identification whenever possible and that the Hispanic-origin question precede the race question on the questionnaires. Since the 1990 census, the Hispanic-origin question has changed very little, but the new instruction for the race question asked respondents to mark one or more races depending upon how individual respondents viewed their own racial identity.

The preliminary guidance on tabulation indicated that Federal statistical agencies should present as much information as possible in their data tables, consistent with maintaining the confidentiality and quality of the data.

The Public Law (P.L.) 94-171 redistricting data file was one of the first configured according to the new rules. The six census race categories are—

- White.

- Black or African American.

- American Indian and Alaska Native.

- Asian.

- Native Hawaiian and Other Pacific Islander.

- Some other race.

In addition to the 6 individual categories, there are also 57 combinations of these categories, or a total of 63 possible responses. Each of these options will be tabulated for total population, total non-Hispanic population, population 18 years of age or older, and non-Hispanic population 18 years or older. Adding in the 14 other population group totals brings the number of cells of data for each unit of census geography to 266. Since there are also 15 check-off response categories and 3 write-in responses with a maximum of 2 entries per write-in, the number of data cells per table could rise to tens of thousands.

The Bureau's new Data Access and Dissemination System (DADS) will allow users tremendous flexibility in accessing these data via the Internet. Data users will be able to specify the tabulations they want for different racial and Hispanic-origin groups, geographical

units, subject matter, and output medium. In addition, different divisions within the Bureau may offer more specialized products with greater detail in terms of context, but less geographic specificity, as funding and staffing allow.

The Bureau's general approach to predefined data products from Census 2000 will be to release fewer products than were published following the 1990 census. The products that are released will have less detailed content and will present characteristics for "major" race categories and total Hispanic groups for "major" geographic units. Some totals for detailed race and Hispanic-origin groups will also be released.

Dr. Tucker of the U.S. Bureau of Labor Statistics (BLS) pointed out that developing the guidelines for implementing the new OMB standards has been under way since September 1997. In August 1997, two areas for further work were identified—meeting user needs and procedural design. The October 1997 *Federal Register* notice identified two sets of data users—

- Those carrying out constitutional or legislative mandates.
- Those monitoring economic and social trends or conducting evaluations.

Three options for data-product layouts for the first group of data users were recommended—

- Full detail, i.e., all cross classifications of racial and ethnic groups.
- Collapsed distribution, tabular display of racial and ethnic groups based on the requirements of data users.
- All-inclusive distribution would include all racial categories respondents indicated; where respondents indicated two or more races, they would be tabulated in all the racial groups indicated.

For the second group of data users, several issues were under consideration—

- Effects of the new standards on control counts and weighting in large national surveys.
- Problem of distinguishing between actual changes in the data and changes due to new standards.
- Calculation of rates, such as poverty rates and birth and death rates.
- Maintenance of confidentiality in table cells with small counts.

To deal with the issues of both types of data users, the working groups were formed. The first of these focuses on issues related to the policy needs of Federal departments and agencies, such as assessing those needs and the types of programs being evaluated, and the services they provide. The second group is a technical group that concentrates on methods for data tabulation. It was split into two subgroups—

- Data quality and related analytical issues.

- Examining ways to separate methodological change from true change in population counts.

The third group focuses on procedural design and is responsible for developing guidelines for data collection. This group is composed of behavioral scientists from the BLS, Census Bureau, National Center for Health Statistics, Government Accounting Office, and other agencies. They are designing questions that take into account the effects of different data collection methods (e.g., telephone, mail, and personal visit) and respondent characteristics (such as language and racial subpopulations). Another area under investigation is administrative record data collection, particularly the recording of race in school records and on birth and death certificates. This group also is examining data collection procedures, such as interviewer training and instructions to both interviewers and respondents.

These work groups expect to produce preliminary reports in June 1998. The goal is to release a set of guidelines for implementing the new OMB standards by the end of 1998.

Dr. Wilson of the Population Association of America (PAA) supported these efforts to improve racial reporting and was pleased that Federal agencies will report all the information available in the way respondents originally reported it. Data users can construct their own classifications.

On the issue of race reporting in administrative records, he expressed concern about the racial categories being used and wanted the Government to encourage institutions to report as much data as possible in as much detail as possible. Given the existence of the new standards, it will be very important to come to some consensus on reporting race in the records of both public and private institutions.

With regard to the reporting of race in data tables, he preferred an inclusive approach in which, for example, individuals who reported themselves as both Black and American Indian would appear in both categories and in a combined Black/American Indian category. Admittedly, this might pose problems for historical analysis. Another distribution he favored would randomly allocate multiracial respondents to one of the race categories. While this type of table would not be comparable to earlier data, it might be useful for analysts.

He also felt uncomfortable with the “other race” category because he did not know what it was supposed to communicate. If the category consisted largely of Hispanics, there should be some way to indicate that. To the extent possible, he would like to see the number of respondents in this category reduced to zero.

Dr. Bell of the American Statistical Association (ASA) said he strongly supported the modifications to the race and ethnic questions and offered three principles for tabulation under the new system—

- It would be a mistake to reclassify to a single race respondents who checked two or more race categories.

- For pie charts and cross tabulations, data users would be most comfortable with results that add to 100 percent.

- It would be preferable not to report results for very small race combinations. While it is difficult to establish exact boundaries, there are reliability and confidentiality problems with small groups. However, in other types of tables (for example, a particular characteristic tabulated by race), totals need not add to 100 percent.

Ms. Gordon pointed out that the OMB will issue the guidelines on tabulating racial and ethnic data and that these will apply to all Federal data-collection efforts. The OMB is concerned with maintaining consistency and comparability in data collection across agency boundaries. The Census Bureau is a member of the advisory group that is advising the OMB on the process of developing the guidelines. Census 2000 will be the first large-scale example of data collected and tabulated under the new guidelines. The Census Bureau is also involved in the second issue, categorizing data collected under the new guidelines in ways that will allow comparisons to past data.

Dr. del Pinal said Dr. Wilson's concept of inclusive distributions was interesting but noted that this categorization would create large numbers of tables. The Census Bureau will have to categorize the data in ways that will permit comparisons across tables. For the purpose of comparison to earlier data, one option would be to combine people who said they were both White and Black into the Black category. He characterized this approach as a minority preference model and wondered how the Bureau should deal with people who report two or more minority races. Another alternative would be to use a smallest group approach, in which individuals reporting two or more minority heritages would be assigned to whichever group was the smallest.

Traditionally, the largest component of the "other race" category has been Hispanics. However, there is a core group that is not Hispanic. One significant group is Cape Verdeans; most of the remaining non-Hispanics give multiracial backgrounds.

On the issue of reporting on small population groups, some have suggested using 0.1 percent or some other arbitrary limit. One problem with this approach is the localization of certain population groups. Nationally, American Indians and Hispanics overlapped by 8.4 percent in 1990; in certain local areas, the overlap jumps to 80 percent. Groups that are small at the national level may loom very large in local jurisdictions. One of the benefits of releasing fully detailed data sets is that researchers can reaggregate population groups based on their own research strategies.

Dr. Tucker pointed out that the members of the technical committee and the procedural design group are examining racial and ethnic reporting in administrative records and evaluating changes in the reporting forms. The data quality analysis group is considering many tabulation options for race and ethnic data. In addition to the data themselves, the group also has to consider the impact of geography on tabular presentation.

Dr. Passel (PAA) pointed out that comparability and continuity will be immense problems with the new standards. Historical comparability can not be achieved but can be approximated. The data that appear under table headings in decennial census tables, for example, are approximately comparable but not exactly the same. Changes in definitions, in processing routines, and in the world at large conspire against comparability. Similarly, the data-collection environment also changes over time. He would like to have the Bureau give researchers some

guidance on ways to proceed from complex data to mutually exclusive groups that add to 100 percent. Dr. Myers (PAA) agreed that data users want tables and charts that add to 100 percent. The Bureau already has some experience in dealing with complex categories, since it has had to confront the overlap between Hispanics and the race categories. One way to resolve this is to make the overlap Hispanic dominant, i.e. maximize the number of Hispanics and subtract people out of other categories. This is a very messy issue, and any solution will have some bias.

Develop Recommendations and Special Interest Activities (AEA)

Dr. Scherer suggested there were three principal areas on which the subgroup might make recommendations—the Bureau's poverty measurement program, the Research Data Centers (RDCs) program, and the research and development (R&D) survey and program. (See Appendix A for the recommendations made and the Bureau's responses.)

Develop Recommendations and Special Interest Activities (AMA)

Dr. Spiro discussed the three alternatives she prepared as her recommendation for marketing Census 2000 products before submitting the recommendations to the American Marketing Association subgroup. Other members decided to forego any discussion and submitted their recommendations in writing. The AMA subgroup commended the Bureau for measuring key points of the marketing "hierarchy of effects" model, i.e., awareness, attitudes, intent, and behavior. The recommendations were based on the questions and topics submitted by the Census Bureau. (See Appendix A for the recommendations made and the Bureau's responses.)

Develop Recommendations and Special Interest Activities (ASA)

Dr. Bell of the American Statistical Association (ASA) reviewed the general topics covered by draft recommendations. (See Appendix A for the recommendations made and the Bureau's responses.)

Develop Recommendations and Special Interest Activities (PAA)

Dr. Klerman asked the members of the Population Association of America subgroup for agenda items for the next Committee meeting. The suggested items included-

- Updates on Census 2000, sampling and dress rehearsal results.
- Presentation and demonstration of software for Post Enumeration Survey and Integrated Coverage Measurement.
- Research status of Census 2000 household file.

Subgroup members brought written drafts of recommendations with them to the session. (See Appendix A for recommendations made and the Bureau's responses.)

Closing Session

Continued committee and staff discussion. Dr. Scherer of the American Economic Association (AEA) subgroup said the last two agenda items are to report out the recommendations from the subgroups, and to make suggestions for the agenda of the fall meeting. Spokespersons for the professional associations reviewed the recommendations prepared or under consideration by their respective subgroups. (See Appendix A for the Committee's recommendations and the Bureau responses.)

Plans and suggested agenda items for next meeting. Dr. Scherer commented that the AEA subgroup has suggested that the next meeting's agenda include sessions on the Bureau's work on the information sector of the economy and on database linkage programs. Dr. Passel of the Population Association of America (PAA) subgroup said the PAA members recommended the meeting include an update on Census 2000 activities, specifically more on the dress rehearsal results and on hiring issues; an update on sampling; a detailed presentation on issues and software for the Post Enumeration Survey (PES) and Integrated Coverage Management (ICM); and a presentation on the household file for Census 2000. Dr. Bell of the American Statistical Association (ASA) added that the ASA subgroup agreed with all the proposed agenda items for the next meeting, and requested a presentation on small-area estimation for poverty.

There were no public comments.

The meeting adjourned at 12:25 p.m.

I hereby certify that the above minutes represent an accurate record of the proceedings of the meeting held on April 23-24, 1998, by the Census Advisory Committee of the Professional Associations.

Frederick M. Scherer, Chairperson
Census Advisory Committee of
Professional Associations

APPENDIX A
Recommendations and Census Bureau Responses

RECOMMENDATIONS OF THE CENSUS ADVISORY COMMITTEE OF PROFESSIONAL ASSOCIATIONS MADE AS A RESULT OF THE MEETING ON APRIL 23-24, 1998

The Census Advisory Committee of Professional Associations made the following recommendations to the Director, Bureau of the Census. Comments showing the response and action taken or to be taken by the Census Bureau accompany each recommendation.

Recommendation 1

Plans for Poverty Measurement

"The Committee commends the Census Bureau for its research in partnership with other Federal agencies to test and implement recommendations from the National Academy of Sciences (NAS) regarding revision of the poverty definition. The proposal to use the Survey of Income and Program Participation (SIPP) as the basis for measuring poverty presents opportunities to improve the measurement of poverty but also creates challenges for subnational analysis and does not rule out use of the Current Population Survey as a valuable auxiliary source of information. The Committee also is concerned with how redefinition of poverty based on SIPP would affect poverty measured in the decennial census, in the American Community Survey, and in the Small Area Income and Poverty Estimates (SAIPE) program. The Committee recommends that:

"1) Medical costs be broken out as a distinguishable component of costs of living for the purpose of defining poverty thresholds and that further research on this matter is needed.

"2) Research is needed on accounting for attrition bias in using successive waves of SIPP, and assuming that attrition bias remains the same from year to year may not always be true.

"3) Further research is needed on the stability of 'family units' (especially the treatment of cohabiting couples) in year-to-year measurements of poverty, appropriate equivalence scales for different family composition, and access of family members to pooled resources.

"4) Special attention be given to defining poverty among the elderly in light of possible reforms to Social Security and Medicare in the future.

"5) Research is needed on geographic cost-of-living variations and the overall estimation of poverty in small areas.

"Overall, the Committee recognizes that the Census Bureau's responsibility is to provide research and data to support the measurement of poverty, but final selection is a separate policy decision."

Census Bureau Response

The Census Bureau is aware of and concerned about differences in poverty estimates produced by SIPP and the Current Population Survey (CPS). Our research is intended to evaluate reasons for the differences. An Office of Management and Budget (OMB) technical work group has been formed to assist the Census Bureau in conducting research on revising the poverty definition in a way that is applicable throughout the federal statistical system.

Medical costs. As suggested by the National Academy of Sciences' panel, we are investigating the feasibility of subtracting medical out-of-pocket costs from disposable income in determining resources. Our current focus is on the feasibility of using imputed medical out-of-pocket costs since there is no practical way of collecting a good measure of these amounts in conjunction with the other detailed information needed for alternative poverty measures. We have included questions in the 1996 panel of SIPP to allow a reasonably accurate statistical match with the Medical Expenditure Panel Survey as an improved method that could be used with SIPP estimates of poverty.

As an alternative to the panel's recommended approach (deducting medical out-of-pocket expenses from income), we are also investigating whether we should instead include work and medical expenses in the thresholds. (Increasing thresholds without changing income has the same effect as reducing income without changing poverty thresholds.) We have discussed with staff at the Bureau of Labor Statistics (BLS) the feasibility of including medical out-of-pocket expenses in their computation of thresholds. Including these expenses in the basic bundle for the reference unit of a two-adult and two-child family and assigning thresholds based only on family size and composition is similar to subtracting an average amount of medical out-of-pocket expenses from the resource side based only on these characteristics.

Finally, we also are working with staff at the Department of Health and Human Services to actively promote research into the construction of a Medical Care Risk Index as proposed by the NAS. Such a procedure might also yield more suitable methods for valuing subsidized health insurance than have been available in the past.

Attrition bias in SIPP. The Census Bureau objective in developing a design for SIPP to be the official source of poverty data is to produce direct survey estimates of poverty with the same reliability levels as currently achieved by the CPS March supplement as well as consistent estimates of change. One of the goals is to eliminate trend biases in the estimates due to respondent attrition. For that reason, prepare SIPP to have NAS panel each year. The three-year rotating panel produces a "steady state" in terms of the amount of attrition in each yearly estimates and addresses this goal.

The Committee has made a valid point that any attrition that may be linked to levels of income implies a bias in the overall level of the poverty estimates. Our intention is to monitor the level of attrition in SIPP and to evaluate its effect on the level of the poverty estimates. Some of this work can be done as data from the current SIPP panel become available. Some of the work must wait until we begin to test the new design and determine the level of attrition in that design. Depending on the outcome of these evaluations, several options are possible. These include the Committee's suggestion for research on adjustments for attrition. They also include changes in the design of the poverty sample and changes in how we handle field procedures that may influence the differential attrition. The Census Bureau regards these evaluations as an essential part of the effort to develop SIPP as the basis for the official poverty measure.

Model-based estimates are an alternative to survey-based estimates but will not be available as quickly. We will continue to investigate such approaches as part of the SAIPE program.

Family units and cohabiting couples. The Census Bureau has initiated research on the stability of "family units" using the measure of migration in the March CPS for persons in family units as officially defined in the current poverty measure and for alternative approaches suggested by the NAS panel. The migration questions measure whether each person lived in the same housing unit at the survey date and one year earlier. We also have reviewed published information from the American Housing Survey to assess the income of persons living together to draw general conclusions on resource pooling to share housing costs. Any working paper developed will be posted on the Census Bureau's poverty measurement web site: htttp://www.census.gov/hhes/ www/povmeas.html.

The elderly. Special attention is being given to the elderly. The technical interagency work group formed by OMB includes two members from the Social Security Administration. They have expressed a wish to be informed of all aspects of the work and to be provided additional measures to fully understand the implications for the elderly of any proposed changes in the measure of poverty.

Geographic cost-of-living variations. The NAS recommended that appropriate agencies conduct research to improve the estimation of geographic cost-of-living differences in housing as well as other components of the poverty budget. This issue is currently being addressed by an interagency work group chaired by OMB. Staff at BLS have taken the lead in this area.

We are mindful that our purpose is to provide information that facilitates careful consideration of the many alternatives and illustrates the possible effects that the choice among the alternatives has on various subgroups of the population and on our understanding of poverty in general. The OMB will ultimately decide whether, and if so how, to modify the current definition of poverty.

Recommendation 2

Corporate Marketing Program

“The essential ingredients of a market plan are presented in this proposal. Our recommendations are refinements that may enhance the outcome given the limited budget.

“Regarding the corporate marketing program we recommend that the target segments be refined and also possibly reduced in number; that consideration be given to the product, price, and distribution elements of the marketing mix as well as promotion; and that the promotion efforts be focused on vehicles most likely to generate sales as opposed to less ambitious goals.”

Census Bureau Response

In keeping with the comments and recommendations made by the AMA Advisory Committee members, the Corporate Marketing Program will be limited to three, more narrowly-focused audiences consisting of college and university librarians, research and testing organizations, and libraries of international organizations. These markets will be reached using the same catalog, although a different cover emphasizing target market benefits will be designed for each. Plans to attend and exhibit at conferences and conventions where no financial commitment or contractual obligation has been made will be canceled and the resulting savings redirected to increase promotional frequency. Simultaneously, the Marketing Services Office will take promotional advantage of exhibits planned and attended by other Bureau divisional staff.

Recommendation 3

Economic Census Update

“Industry analysts at the Census Bureau should to the maximum degree possible track new developments in the industries for which they are responsible using such bibliographic sources as the Predicasts Funk & Scott Index and the Wall Street Journal Index.”

Census Bureau Response

We support the Committee’s suggestion to access bibliographic sources to track new developments in industries. For years we have utilized paper sources, such as Moodys, Dun and Bradstreet, and the Thomas Register in their analysis of industry data. Recently, we have begun to make use of bibliographic resources on the World Wide Web. For example, the intranet site maintained by the Economic Directorate provides links to a variety of sites including "Hoover's Online," "Industrylink," and the Securities and Exchange Commission's "Edgar." We will explore acquisition of site licenses for staff to access the Wall Street Journal Index and the Predicasts Funk & Scott Index.

Recommendation 4

How can the Census Bureau get Consistent and Useful Feedback from its Customers?

"We applaud your organized approach to generating additional customer feedback. This program is so multi-faceted that it will be useful to schedule periodic reviews for program modifications, additions and deletions.

"We endorse the expansion of customer feedback into electronic media and agree that careful controls are necessary in live on-line forums. Approaches to consider offering more control are those with delayed posting of messages and questions such as cybertalk town meetings (questions submitted in advance), closed bulletin boards and listserver forums. Registration/password requirements for live chats may add control.

"We also support Bureau-sponsored conferences and forums as a way to interact face-to-face with customers as well as facilitating user interaction."

Census Bureau Response

Based on Committee comments regarding access to consistent and useful feedback from its customers, the Bureau plans to:

- Identify opportunities for future user conferences to build relationships and to allow for better networking.
- Proceed with its proposal for a web-based Customer Information Exchange Page.
- Investigate more fully options to the "live" chat room proposal, including delayed response mechanisms, listservers, and bulletin boards.
- Look into warranty/registration cards with product distributions to build more customer information into its databases.
- Find incentives to encourage customer services representatives to identify and to regularly report customer-generated issues.

We will keep the Committee informed of our activities and progress.

Recommendation 5

Sampling and Estimating in Census 2000 and the Dress Rehearsal

"We commend the Census Bureau on the progress it has made on sample design and estimation decisions. In particular, the Census Bureau has produced creative solutions for dealing with late mail returns and large Integrated Covered Measurement (ICM) block clusters. However, remaining decisions need to be made very soon so that they can receive adequate outside review.

"With respect to whether to use sampling in Census 2000, the crucial question is not 'Is sampling perfect?' but 'Is sampling, in practice, better than an attempt at conventional enumeration, in practice?' Both delegations recognize a consensus of the demographic and statistical communities that the appropriate use of statistical sampling for nonresponse in Census 2000 is likely in practice to yield better results than conventional enumeration. Sampling is likely to increase accuracy, decrease costs, and facilitate implementing the census. This was and continues to be the position of the NAS panel. It is the position of the ASA and PAA delegations to the Census Advisory Committee of Professional Associations.

"Our support for the basic principle of sampling should not be understood as a blanket endorsement of the specific procedures currently proposed by the Census Bureau for implementing a 2000 Census incorporating sampling. While the Census Bureau has specified its general approach to Census 2000 procedures, many details have not yet been released for external review. Thorough internal and external review of explicit plans (down to the level of operating computer programs and sample output using test data) will improve procedures, reduce the change for unexpected problems, and minimize concerns about manipulation. Of course, final procedures may need to deviate from detailed plans to accommodate unanticipated circumstances.

"In finalizing plans, the Census Bureau should evaluate modifications that might improve the accuracy of the population counts, cut costs, or facilitate successful completion of all procedures on schedule. Careful study of these issues by Census Bureau staff, early release of precise plans, and their review by the general public, the NAS panel, the advisory committees, and other experts will best ensure the proper decisions leading to a successful census in 2000."

Census Bureau Response

As details of our general approach to Census 2000 procedures become available, we have been seeking their internal and external review and will continue to do so. In addition to the CNSTAT (Committee on National Statistics, National Academy of Sciences) panels and Census Advisory Committees, we have established three Quality Review Boards to review our detailed procedures for estimation of nonresponse, for estimation of coverage measurement, and for unduplication. The reviews from these different groups will help us improve procedures, reduce the chance of unexpected problems, and minimize concerns about manipulation.
The Census Bureau has started processes to review the Census 2000 plans. In January 1998, several Census Design Review meetings were held with various teams in the Census Bureau Decennial Management Division and Decennial Statistical Studies Division to improve the

quality of the population counts and to ensure the successful completion of procedures on schedule. We have identified some issues and are currently exploring them.

Recommendation 6

Innovation and Research and Development (R&D) Statistics

"1) The Census Bureau and the NSF should investigate carefully a highly desirable change in R&D reporting to a business unit basis. The precise set of business unit reporting categories should be worked out to achieve the optimal balance between feasibility and information content. One option is to add questions to MA-1000 asking (1) whether R&D is performed in the reporting establishment; (2) if so, how many FTE employee years were allocated; and (3) (lower priority) the fraction devoted to product and process R&D, respectively. Alternatively, these data might be obtained as adjuncts to the occasional Survey of Manufacturing Technology, or, as a third alternative, from corporate reporting offices. The Committee believes that improvement along this line is so important that mandatory reporting should be pursued. The choice between approaches should take into account among other things the fraction of total corporate R&D carried out at the manufacturing and service operating establishment levels. When business unit data are elicited, a separate question should determine the amount of R&D conducted at the corporate level or at 'not classifiable' industry sectors.

"2) A paramount consideration in redesigning R&D data collection protocols should be maximizing opportunities to link the data to other microdata bases.

"3) The Census Bureau and the NSF should study the problems of R&D reporting by multinational enterprises and develop new, clearer reporting guidelines.

"4) In conjunction with the NSF, the Bureau of Labor Statistics, and the Bureau of Economic Analysis, Census should develop and maintain explicit input cost deflators for industrial R&D expenditures.

"5) The Census Bureau and the NSF should study the feasibility, perhaps using private sector data such as the Directory of Industrial Research and Development Laboratories, of breaking down at least every four years R&D efforts more precisely by Standard Metropolitan Statistical Areas. Such data are expected to have considerable value in tracing R&D spillovers.

"6) Assuming that industrial R&D efforts can be estimated at the business unit level, efforts should be undertaken by the Department of Commerce to have the Patent Office routinely code each domestic patent by industry of origin using the same industrial classification, and code the likely uses of patented inventions, both domestic and foreign, by using industries. A link between R&D origin and patent use data effected by the Center for Economic Studies (CES) is expected to have considerable value in tracking market-mediated productivity impacts.

"7) The Census Bureau and the NSF should explore the feasibility of surveying the output of research, including compiling data on numbers of product and process innovations and the fraction of receipts for which the new products or processes account.

"8) The Census Bureau and the NSF should explore the question of whether there are ways, especially for small firms, of measuring R&D efforts that accompany routine activity.

"9) The AEA Committee members urge the Census Bureau to investigate the best survey instrument for eliciting information on R&D physical capital investments."

Census Bureau Response

In response to the Committee's strong recommendations to collect data at finer industry and geographic area detail on R&D and innovation activity, we intend to work with the National Science Foundation (NSF) to determine if business unit, or other source of detailed end-use industry information, and geographic data can be collected effectively in the Industrial Research and Development Survey. We also will explore the use of alternative survey programs and external information to provide the means for collecting or allocating data by end-use industry or geography.

On the issue of mandatory reporting, we believe introduction of such a requirement would have to be accompanied by some reduction in the data requested from respondents to avoid a backlash from the reporting community. We will explore the possibility with NSF.

As we discuss changes to the Industrial Research and Development Survey with NSF and investigate alternatives for collection of R&D and innovation statistics, we will keep in mind the needs for micro-level data analysis. This includes ways to improve analysis between R&D and innovation data sets and other Census Bureau establishment and enterprise data sets. We will involve the Center for Economic Studies in this effort.

We will continue to work with the NSF on improving the guidelines for reporting foreign R&D in the Industrial Research and Development Survey.

We appreciate the Committee's interest in developing cost deflators for industrial R&D expenditures. While the Census Bureau usually does not typically develop cost deflators, we will advise the Bureau of Economic Analysis and the Bureau of Labor Statistics of the Committee's interest for consideration in their ongoing index programs. We also will accommodate these agencies' requests for data to develop deflators to the extent possible.

In response to the Committee's suggestion to use private sector data, we have attempted to use the Directory of Industrial Research and Development Laboratories to allocate reported R&D expenditures by state when the respondent has not provided the information. We find that the Directory is an excellent source of addresses and contacts for R&D labs. However, information about the size of the labs in terms of employment, number of scientists, R&D spending, or R&D budgets to use to allocate R&D expenditures is inconsistent and frequently not given. In some instances, we could not find listings in the Directory of companies reporting in the R&D survey. We will continue to review other sources of information that might be useful for allocating data by geographic area.

We will contact the Department of Commerce (DOC) Patent and Trademark Office (PTO) to explore what information is available currently and what information could be obtained to relate patent information to R&D activity by industry. A recent DOC policy agenda encourages the PTO, in conjunction with the Economics and Statistics Administration, to explore the use of patent and trademark data as a partial barometer of future economic activities.

We agree with the intent of the Committee's recommendation, but we believe that measures of output of R&D would be more appropriately collected in surveys other than the Industrial Research and Development Survey, which collects innovation inputs. We will explore the practicality of collecting product and process innovations through existing surveys if a special survey cannot be undertaken.

We will continue to work with the NSF to improve survey instruments to capture R&D performed in small companies. As part of our survey review, we routinely use external sources of information, such as the CompuStat service, Securities and Exchange 10K reports, and lists of companies with Federal contracts for R&D, to identify R&D performing companies and to validate the level of R&D expenditures.

Several years ago, we asked respondents about their ability to report capital investments in R&D. A large percentage told us that they had separate records for R&D capital expenditures. However, the records were maintained by offices in locations other than those reporting expenditures. Other companies could not distinguish R&D procurement from other capital expenditures without reviewing all procurement. Still others reported that the property may be shared and determining the R&D share was not straightforward.

Recently, we asked respondents if they could report depreciation of R&D assets. Response was more positive. The recent positive responses to the inquiry on depreciation may be an indication that recordkeeping for capital expenditures has changed. At the next opportunity, we will propose asking about current recordkeeping practices for R&D capital expenditures to determine if access to these records has improved. Based on our findings, we can address the issue of the most effective vehicle for capturing this information.

<u>Recommendation 7</u>

Evaluating the Dress Rehearsal and Census 2000

"We commend the Census Bureau on plans for evaluating the dress rehearsal. We recommend that data be collected during the 2000 census operation that will allow modeling of the relationship between interviewer characteristics and the quality of their work (number and type of last resort data, item nonresponse rate, length of employment, number of interviews completed, etc.).

"We recommend that the Census Bureau propose a more formal program of research on reporting errors in the decennial census. This might include a small reinterview program for the long and short form and an analysis of the reinterview data for the ICM. The purpose of the research would be to increase our understanding of response variance and the mechanisms producing missing data."

Census Bureau Response

We will begin to plan the Census 2000 evaluation program this summer. The types of measures you suggest collecting and the research on reporting errors are in agreement with the types of evaluations we will be planning. We will take your recommendations into consideration as we define the Census 2000 evaluation program.

Recommendation 8

Evaluating the Marketing Strategy for the Dress Rehearsal and Census 2000

"The dress rehearsal offers a very limited opportunity for evaluating the effectiveness of the marketing activities for improving response to the census. The impact of these activities on awareness is confounded with each other, as well as the delivery of the census form and unpaid media exposure. Similarly, effects of the additional activities in South Carolina will be confounded with the many other differences between Sacramento and South Carolina. Furthermore, demonstrating a correlation between awareness and response would not imply that any marketing effects on awareness would spill over to response, since such a correlation could be due to other factors. The linking of survey responses to information about whether a household responded in the dress rehearsal will provide valuable data. The analysis of these data should focus on identifying factors that may affect response in 2000.

"In order to directly measure the effects of advertising for Census 2000 independent of the effects of the mailed census pieces, we recommend that the Census Bureau collect an additional wave of data from respondents after paid advertising has begun but before any Census 2000 pieces have been mailed."

Census Bureau Response

The Census Bureau recognizes the limitations of evaluating the paid advertising activities in the 1998 Dress Rehearsal sites. Our pre- and post-wave survey design will not clearly separate the influences of paid media from unpaid media (for example, Census Partnership activities and the receipt of census pre-notices, forms, and postcards).

We also agree that the evaluation of the "Innovative and Aggressive" (I&A) media campaign in South Carolina is extremely difficult due to market and demographic differences between South Carolina and Sacramento. We did, however, add two specific questions to the post-wave questionnaire to try to assess certain elements of the I&A efforts.

We agree that the linking of survey responses to information about census method of response (mail or not) will provide valuable data. For this reason, we have proposed to match post-wave cases to information on census data files. This will provide a validation of self-responses about mail-back behavior and improve the quality of our dependent variable. According to the subcommittee's recommendation, we will concentrate our logistic regression models of variables that may affect response in 2000 (for example, source of media exposure, recall of media message, knowledge of census uses, and so forth).

The Census Bureau agrees with the recommendation of collecting an additional wave after paid advertising has begun but before census mailing pieces are delivered. Due to resource and timing constraints, we were unable to implement such a design for the 1998 Dress Rehearsal evaluation. However, for Census 2000 we will try to build an interim wave into the research design.

Recommendation 9

Pricing Data Products through Data Access and Dissemination System (DADS)

"The Census Bureau seems to have conflicting pricing objectives. One is to achieve the widest possible dissemination possible and another is to recoup distribution costs. We recommend that the Census Bureau assess its strategic priorities and decide which objective is more important.

"Specific recommendations:

"1) Conduct an analysis of competitors' products, services distribution, and pricing.

"2) If Census Bureau products cannot be differentiated from those of other providers, you must price them at the same level as other providers or lower if you wish to achieve the widest distribution possible.

"3) If some of your products and services can be differentiated—that is, they are perceived by the consumer as adding value over those products of other providers—then you should price the products according to the value of the product as perceived by the consumer or lower if you wish to achieve wider dissemination. The perceptions of consumer value must be determined through research.

"Whether the Census Bureau should seek to be able to copyright their products remains an issue. Basic products should be provided to individuals free of charge. But no one should be allowed to copy or distribute this information or any altered form of the information for resale without the permission of the Census Bureau. Additional products provided by the Census Bureau should be priced according to the value as perceived by the end user.

"Another possibility is that the Census Bureau should not continue to develop or disseminate any products other than the most basic information to the final consumer. The Census Bureau might partner with external providers who will then sell additional products to the final consumer. Some providers are already offering competitive products at significantly lower prices than those offered or proposed by the Census Bureau.

"We support the general principle that special-purpose tabulations should not be performed unless they cover their fully allocated costs. It is reasonable to charge some portion of the overhead for such tabulations. But attempting to recover all (or even a substantial fraction) of the broadly defined "dissemination costs" through charges on the distribution of off-the-shelf census products is doomed to failure, for the secondary market will buy the data and resell it, often with only minimal repackaging.

"We recommend that the Census Bureau convene a meeting devoted to the issue of data products and dissemination. The meeting should include Census Bureau staff from different divisions involved in product creation and dissemination, together with very experienced data users and intermediaries from outside the Census Bureau."

Census Bureau Response

We are reviewing the costs associated with product production and dissemination and will use the information to examine the prices we are charging for our CD-ROMS. We also are reviewing the prices charged by competitors. Throughout this exercise, we will be trying to balance the often conflicting goals of garnering and maintaining a broad base of support for our data collection efforts, through providing basic data at little or no cost, while recovery of direct cost is important because of the uncertainty of appropriated funds to fully support our future dissemination efforts. We also will attempt to use a pricing mechanism to reduce the impact on the DADS of large, complex data requests.

While we agree with the Committees' suggestions that copyrighting our information would provide greater opportunities for partnering, pricing, and sales, we have reviewed the steps necessary, including legislative changes, and have concluded that obtaining copyright authority is not a feasible option at this time.

Once we have reached internal consensus on our Census 2000 products proposal, we plan to communicate these plans to data users in a variety of venues, including meetings such as those suggested by the Committee.

Recommendation 10

General Edit and Imputation Research

"The Committee commends the Census Bureau for conducting leading-edge research on general edit and imputation methods and on algorithms that dramatically increase the speed of programs that implement these methods. The Committee encourages the Census Bureau to continue developing general editing and imputation software. We also encourage the Census Bureau to continue the promising research into modeling approaches for imputation. We encourage the Census Bureau to test these methods on additional data sets to evaluate the accuracy of the editing and imputation procedures."

Census Bureau Response

We intend to further develop methods and software for general editing and imputation. For our economic surveys, we will create additional enhancements to Structure Programs for Economic Editing and Referrals. We hope to apply the DISCRETE edit system to a variety of demographic surveys. We intend to develop a generalized imputation software package that could be applied in a moderate number of surveys and to create the training methods to assure its success. In all situations, we will continue to develop model-based imputation methods and test them on different surveys.

Recommendation 11

Chief Economist Update

"The Committee members from the AEA appreciate being updated on the status of the Center for Economic Studies (CES). We enthusiastically commend both the opening of new regional research centers and the expansion of CES data bases. The possibility of transforming CES data into public use files, taking into account appropriate confidentiality considerations, should be pursued.

"We especially commend the effort to link demographic and economic census materials and to create longitudinal employer-employee matched files."

Census Bureau Response

The Census Bureau is committed to the expansion of research data centers and data sets that can be used at these centers. We are excited about the possibility of combining establishment and demographic data and, in particular, longitudinal matched employer-employee data. We have some pilot projects underway to study feasibility and to develop protocol for database development and access. Development and access must be pursued in a manner that maintains protection for privacy and confidentiality. We cosponsored a major conference in May 1998 to help address these issues.

We appreciate the demand for public use data that might emerge from the database development work at the CES. We will take this demand into consideration as we proceed with the efforts described above.

Recommendation 12

How Should We Promote Confidentiality in the Decennial Census?

"We commend the Bureau in raising this issue for thorough review and resolution. We agree that confidentiality and privacy concerns will directly effect response rates. The question relates to the role of advertising and other communication channels which could or could not address this key issue:

"1) Since the advertising goal is to increase response rate, and a number of studies link that the perception of confidentiality is correlated with response rate, advertising should address confidentiality.

"2) Confidentiality should be addressed in all channels of communication. We agree that partnerships and other local community leaders will be especially effective and credible—yet this should not be an 'either/or' debate, a re-assurance on privacy needs to be included in all communication.

"3) We agree with Y & R's position that a benefits positioning 'what's in it for me' should be the primary focus and message—yet a privacy re-assurance, i.e., 'By law, your census information will be kept (private? confidential?)' is not incompatible with the benefit emphasis—and can be included as a key secondary message.

"4) We feel there is too much reliance on qualitative focus group research (questions of lack of reliability/validity/benchmarking)—we recommend using the dress rehearsal to prove and pin down the issue—dissect the 'non response' sample, analyze pre/post survey.

"5) We feel the issue of confidentiality and privacy needs to be thoroughly examined and addressed during the next 12 months.**"**

Census Bureau Response

The Census Bureau welcomes the Committee's recommendations regarding confidentiality in the Decennial Census.

We realize that concerns about confidentiality are a major barrier to participation, particularly among some key target audiences. However, we are not sure that offering assurances in advertising is the answer. For those concerned about this issue, is advertising a credible source for such assurances? For those not currently concerned about privacy, does mentioning it in advertising raise fears? The best way to resolve these questions is to conduct quantitative research among a cross-section of the population, including both those more likely and less likely to participate. This research will enable us to test the impact of communicating such a message in advertising—does it resolve concerns among those for whom it is more of an issue and does it have any effect on those who are not currently concerned with it?

The advertising for Census 2000 Dress Rehearsal carried some targeted messages on confidentiality. The degree to which confidentiality will be incorporated into advertising for the

general public will depend on our continuing research. All partnership materials that are under development at the present time will carry confidentiality messages.

We strongly and unanimously recommend a pre-test/copy test of advertising, using one of the leading advertising research companies (McCollum-Spielman, ASI, ARS, Mapes and Ross, and so forth). Testing can be at the "rough" stage (with multiple executions) or "finished" stage (fewer executions).

We plan to use more quantitative research as we work toward 2000. We believe qualitative and quantitative research are both valuable, providing different types of information. In developing advertising for the Dress Rehearsal, our time was limited and, therefore, we had to limit the research we were able to conduct. To refine the creative work, we conducted the "Roper" quantitative research to confirm the strategy and focus groups among target audiences. For 2000, we plan to use qualitative research to develop hypotheses and gain early feedback on creative work. We will use quantitative research to answer strategic questions and test copy.

The only way to resolve the questions regarding confidentiality will be to examine this issue during the next 12 months.

Recommendation 13

Implementing the OMB Proposal on Tabulation of Race and Ethnicity Data in Dress Rehearsal Tabulations

"The Committee commends the Census Bureau for its careful and thorough research on collection of race and Hispanic origin data. We recognize the complexity of data collection and, especially, tabulation issues for these items. In particular, the demands for fully detailed data presentations and for easily usable categorizations are inherently contradictory. It is our opinion that most data users, including the Census Bureau, will require one-way and cross-tabulations that use a relatively small number of mutually exclusive and exhaustive race/Hispanic categories—that is, the categories need to total 100 percent. Further, the Census Bureau, based on its research, experience, and access to fully-detailed data, will be in position to provide users with guidance on the feasibility and desirability of alternative categorization schemes.

"We recommend that:

"1) The Census Bureau provide one set of tabulations that does not reclassify persons who choose multiple races into one of the traditional race categories. In the interests of maintaining confidentiality and reliability, these tabulations, may, however, require that some small cells be collapsed.

"2) While OMB may provide some guidelines for tabulations and collapsing categories, the Census Bureau should also include options for various analytic purposes. We recognize, however, that some users will need to prepare their own alternatives.

"3) The Census Bureau should also consider what alternative tabulations may be needed to produce data that are comparable to previous censuses, current surveys such as the Current Population Survey, and other data systems such as vital statistics.

"4) We recognize that there are likely to be objections to any particular choice for collapsed data, but we also recognize that some choices will have to be made to make the data manageable. To this end, we recommend that alternatives be made available to users and that the Census Bureau guide their own choices, where possible, by the results of research conducted on preferences given alternative categories."

Census Bureau Response

The Census Bureau thanks the Committee for its recommendations, including the recommendations to collapse race/Hispanic categories into mutually exclusive categories and to guide this collapsing by research on respondents' preferences given a single choice on race. We will pass these recommendations on to the OMB Interagency Tabulation Working Group for their consideration. We will keep the Committee informed about the progress of OMB's process to develop tabulation guidelines. We do plan to provide alternative categorizations in the DADS, but we will work within OMB's forthcoming guidelines.

APPENDIX B
AGENDA

April 20, 1998

Agenda for the April 23-24, 1998, Meeting of the Census Advisory Committee of Professional Associations*

Embassy Suites Hotel
1250 22nd Street, NW
Washington, DC 20036

Thursday, April 23

PLENARY (9:00 - 10:00 a.m.)

Joint Session
Introductory Remarks
James Holmes
Acting Director
Ballroom

PLENARY (10:00 - 10:15 a.m.)

Census Bureau Responses to Committee
Recommendations/Report on October 1997
Meeting, *Frederic Scherer,* Chairperson
Ballroom

BREAK (10:15 - 10:30 a.m.)

AEA, ASA, PAA (10:30 - 11:30 a.m.)

The Census Bureau's Plans for Poverty Measurement
Dan Weinberg, Chief, Housing and Household Economic Statistics Division
Chair: PAA
Ballroom

AMA (10:30 - 11:30 a.m.)

Are We on the Right Track with the Corporate Marketing Program?
George Selby and Josephine Ruffin, Marketing Services Office
Wine Room

AEA (11:30 a.m. - 12:00 p.m.)

Economic Census Update
John Govoni, Economic Planning and Coordination Division
Delegate Room

AMA (11:30 a.m. - 12:30 p.m.)

How Can the Census Bureau Get Consistent and Useful Feedback from its Customers?
Joanne Dickinson, Marketing Services Office
Wine Room

PAA, ASA (11:30 a.m. - 11:45 a.m.)

Overview of the Census 2000 Dress Rehearsal
John Thompson, Associate Director for Decennial Census
Chair: PAA
Ballroom
PAA, ASA (11:45 a.m. - 12:30 p.m.)

*All sessions include question/answer.

Sampling and Estimation in Census 2000 and the Dress Rehearsal
Howard Hogan, Acting Chief, Decennial Statistical Studies Division
Chair: PAA
Ballroom

LUNCH (12:00 - 1: 30 p.m.)

AEA (1:00 - 1:30 p.m.)

Overview of Indicators of Innovation and Technology
John Haltiwanger, Chief Economist
Ron Cooper, Consulting Economist
Delegate Room

AEA (1:30 - 2:15 p.m.)

Panel Discussion: The National Science
Foundation Research and Development Survey
Moderator*: Frederic Scherer*, Harvard University
Delegate Room

AMA (1:45 - 2:45 p.m.)

Demonstration of the Latest DADS Prototype
Marilyn Moore, Sandy Rowland, DADS
Wine Room

ASA, PAA (1:45 - 2:45 p.m.)

How Do We Evaluate the Dress Rehearsal and Census 2000?
Ruth Ann Killion, Office of the Director
Chair: ASA
Ballroom

AEA (2:30 - 5:00 p.m.)

Panel Discussion: Where Should We Go from Here?
Moderator: *Frederick Knickerbocker*, Associate Director for Economic Programs
Delegate Room

BREAK (2:45 - 3:00 p.m.)

AMA, ASA, PAA (3:00 - 4:00 p.m.)

How Do We Evaluate the Marketing Strategy for the Dress Rehearsal and Census 2000?
Nancy Bates, Office of the Director
Chair: AMA
Ballroom

AMA, PAA (4:00 - 5:00 p.m.)

How Should the Census Bureau Price Data Products through DADS?
John Kavaliunas, Marketing Services Office
Chair: AMA
Wine Room

ASA (4:00 - 5:00 p.m.)

General Edit and Imputation Research
William Winkler, Statistical Research Division
Ballroom

ADJOURN

AEA (**9:00 - 10:00 a.m.)**

Chief Economist Updates
*John Haltiwanger***,** Chief Economist
Delegate Room

AMA (**9:00 - 10:00 a.m.)**

How Should We Promote Confidentiality in the Decennial Census?
Judith Waldrop, *Jennifer Marks*, Census 2000 Publicity Office
Gerald Gates, Policy Office
Wine Room

ASA, PAA (9:00 - 10:00 a.m.)

How Will the OMB Proposal on Tabulation of Race and Ethnicity Data be Implemented in Dress Rehearsal Tabulations?
Jorge del Pinal, Population Division
Clyde Tucker, Bureau of Labor Statistics
Chair: PAA
Ballroom

BREAK (10:00 a.m.)

AEA (10:15 - 11:45 a.m.)

Develop Recommendations and Special Interest Activities
Delegate Room

AMA (10:15 - 11:45 a.m.)

Develop Recommendations and Special Interest Activities
Wine Room

ASA (10:15 - 11:45 a.m.)

Develop Recommendations and Special Interest Activities
Boardroom Suite 233

PAA (10:15 - 11:45 a.m.)

Develop Recommendations and Special Interest Activities
Ballroom

CLOSING SESSION (11:45 a.m.)

Continued Committee and Staff Discussion Plans and Suggested Agenda Topics for Next Meeting
Ballroom

ADJOURN (12:15 p.m.)

APPENDIX C
Bureau Personnel Present

Director's Office

James Holmes, Acting Director
Paula Schneider, Principal Associate Director for Programs
Jennifer Marks, Special Assistant to Associate Director for Communications
Nampeo McKenney, Senior Research and Technical Advisor
Frederick T. Knickerbocker, Associate Director for Economic Programs
John Ostenso, Special Assistant
Thomas L. Mesenbourg, Assistant Director for Economic Programs
*Paula Muroff, Special Assistant
Susan Cardiskey
John H. Thompson, Associate Director for Decennial Census
Nancy M. Gordon, Associate Director for Demographic Programs
Campbell Gibson, Demographic Advisor
Cynthia Z. F. Clark, Associate Director for Methodology and Standards
Robert E. Fay, Senior Mathematical Statistician
Ruth Ann Killion

Administrative and Customer Services Division

Kathy Italiano
Kathy Maney
Brenda Williams

Agriculture and Financial Statistics Division

Ewen Wilson, Chief
Ruth Runyan

Census 2000 Publicity Office

Kenneth Meyer, Chief
Mauro Cooper
Judith Waldrop

Chief Economist Office

John Haltiwanger, Chief Economist
Ronald Cooper
Timothy Dunne

Data Access and Dissemination Systems Staff

Marilyn Moore
Sandra Rowland

Decennial Statistical Studies Division

Howard Hogan, Acting Chief
Raj Singh, Assistant Chief

Decennial Management Division

Ramala Basu

Demographic Statistical Methods Division

Larry Cohoon, Assistant Chief

Economic Planning and Coordination Division

John P. Govoni, Chief
Paul Zeisset, Special Assistant

Economic Statistical Methods and Programming Division

Charles P. Pautler, Jr., Chief

Housing and Household Economic Statistics Division

Daniel H. Weinberg, Chief
Donald Hernandez
*Larry Long
'Betty Pittman
Kathy Short

Manufacturing and Construction Division

Stephen Andrews, Assistant Chief
Brian Greenberg, Assistant Chief
Elinor J. Champion, Chief, Special Studies Branch
David J. Gromos
Ron Taylor

Marketing Services Office

*W. Donald Wynegar, Chief
Les Solomon, Chief, Customer Services Branch
Ann Berry
Barbara Garner
Joanne Dickerson

John Kavaliunas
Josephine Ruffin
George Selby
Phil Thompson
Elaine Quesinberry

Policy Office

Gerald Gates, Senior Statistician, Administrative Records
Nick Birnbaum
George Gatewood
Jason Gauthier
Michael A. Hovland
David M. Pemberton
Kathleen Styles

Population Division
Jorge del Pinal, Assistant Chief
Rhonda G. Carney

Statistical Research Division

Tommy Wright, Chief
William Winkler, Principal Researcher
Nancy Bates
Hazel Beaton
Alice Bell
Ann Vacca

APPENDIX D
Membership Lists

APRIL 1998

MEMBERSHIP LIST
CENSUS ADVISORY COMMITTEE OF PROFESSIONAL ASSOCIATIONS
(AMERICAN ECONOMIC ASSOCIATION)

Dr. Ernst R. Berndt (Co-Chair)
E52-452
Massachusetts Institute of Technology
50 Memorial Drive
Cambridge, MA 02139
Term Expiration : 12/31/98
Phone: (617) 253-2665
FAX: (617) 258-6855
Email: erberndt@mit.edu

Dr. Roger R. Betancourt
Department of Economics
University of Maryland
College Park, MD 20742
Term Expiration: 12/31/99
Phone: (301) 405-3479
FAX: (301) 405-3542
Email: betancourt@econ.umd.edu

Dr. Lynn E. Browne
Research Department
Federal Reserve Bank
T-8 600 Atlantic Avenue
Boston, MA 02106
Term Expiration: 12/31/00
Phone: (617) 973-3091
FAX: (617) 973-3957
Email: lynn.browne@bos.frb.org

Dr. Michael Gort
State University of New York-Buffalo
Economic Department
415 Fronczak Hall, North Campus
Buffalo, NY 14260
Term Expiration: 12/31/00
Phone: (716) 839-9776
FAX: (716) 645-2127
Email: gort@acsu.buffalo.edu

Dr. Lee Lillard
The Rand Corporation
1700 Main Steet
Santa Monica, CA 90401
Term Expiration: 12/31/00
Phone: (310) 393-0411 x6535
FAX: (310) 393-4818
Email: Lee_Lillard@rand.org

Dr. Rebecca A. Maynard
Mathematica Policy Research
3700 Walnut Street
Philadelphia, PA 19104-6216
Term Expiration: 12/31/00
Phone: (215) 898-3558
FAX: (215) 573-2241
Email: rebeccam@gse.upenn.edu

Dr. Ariel Pakes
37 Hillhouse
Yale University
New Haven, CT 06520
Term Expiration: 12/31/98
Phone: (203) 432-3550
FAX: (203) 432-6323
Email: ariel@econ.yale.edu

Dr. Frederic M. Scherer (Co-Chair)
Harvard University
Kennedy School of Government
Cambridge, MA 02138
Term Expiration: 12/31/99
Phone: 617-495-9510
FAX: 617-496-0063
Email: Mike_Scherer@harvard.edu

Dr. Robert Willis
3254 ISR
University of Michigan
P.O. Box 1248
Ann Arbor, MI 48106-1248
Term Expiration: 12/31/00
Phone: (313) 936-7261
FAX: (313) 647-1186
Email: rjwillis@isr.umich.edu

Committee Liaison: Ms. Paula Muroff
Room 3061, Bldg. 3
Phone: (301) 457-2846
FAX: (301) 457-2908
Email: Paula.N.Muroff@ccmail.census.gov

APRIL 1998

MEMBERSHIP LIST
CENSUS ADVISORY COMMITTEE OF PROFESSIONAL ASSOCIATIONS
(AMERICAN MARKETING ASSOCIATION)

Mr. Tony Adams (Chair)
Adams Marketing Services
921 Ivycroft Road
Wayne, PA 19087
Term Expiration: 6/30/99
Phone: (610) 688-2666
FAX: (610) 688-2666

Ms. Laurie Ashcraft
President
Ashcraft Research, Inc.
625 N. Michigan Ave., Suite 500
Chicago, IL 60611
Term Expiration: 6/30/98
Phone: (312) 751-5420
FAX: (312) 649-1646
Email: ASHRESINC@aol.com

Dr. Michael Etzel
Department of Marketing
University of Notre Dame
College of Business
Notre Dame, IN 46556
Term Expiration: 6/30/00
Phone: (219) 631-5925
FAX: (219) 631-5255
Email: michael.j. etzel.1@nd.edu

Ms. Beth Fischer
President/Owner
Twin City Interviewing Service
3225 Hennepin Avenue S.
Minneapolis, MN 55408
Term Expiration: 6/30/98
Phone: (612) 823-6214
FAX: (612) 823-6215
Email: tcifischer@sprintmail.com

Ms. Katherine Jocz
Director of Research Management
Marketing Science Institute
1000 Massachusetts Avenue
Cambridge, MA 02138
Term Expiration: 6/30/98
Phone: (617) 491-2060
FAX: (617) 491-2065
Email: kjocz@msi. org

Mr. Arthur Redmond
Managing Director Consumer Insights
Citibank, N.A.
One Court Square
Long Island City, NY 11120
Term Expiration: 6/30/00
Phone: (718) 248-3949
FAX: (718) 248-0962
Email: art.redmond@citicorp.com

Ms. Debra Semans
Vice President
Market Development/Expense Management
Carlson Wagonlit Travel
5897 Windward Parkway, Suite 100
Alpharetta, GA 30202
Term Expiration: 6/30/99
Phone: (770) 753-7703
FAX: (770) 753-7755

Dr. Rosann Spiro
Professor
Indiana University
Department of Marketing
School of Business
Bloomington, IN 47405
Term Expiration: 6/30/00
Phone: (812) 855-8878
FAX: (812) 855-6440
Email: spiro@indiana.edu

Dr. David Stewart
Marketing Professor
University of Southern California
School of Business Administration
Los Angeles, CA 90089-1421
Term Expiration: 6/30/98
Phone: (213) 740-5037
FAX: (213) 740-7828
Email: dstewart@sba.usc.edu

Committee Liaison: Mr. W. Don Wynegar
Room 3023, Bldg. 3
Phone: (301) 457-2155
FAX: (301) 457-2778
Email: dwynegar@census.gov

APRIL 1998

MEMBERSHIP LIST
CENSUS ADVISORY COMMITTEE OF PROFESSIONAL ASSOCIATIONS
(AMERICAN STATISTICAL ASSOCIATION)

Dr. Robert Bell (Chair)
Senior Statistician
The Rand Corporation
1700 Main Street
Santa Monica, CA 90401
Term Expiration: 12/31/00
Phone: (310) 393-0411, x6672
FAX: (310) 451-7004
Email: robert_bell@rand.org

Dr. David A. Binder
Statistics Canada
11-F R.H. Coats Building
Tunney's Pasture
Ottawa, Ontario K1A 0T6
CANADA
Term Expiration: 12/31/00
Phone: (613) 951-0980
FAX: (613) 951-1462
Email: binddav@statcan.ca

Mr. Joseph Garrett
Vice President and Deputy Director
Mathematica Policy Research
600 Maryland Ave., S.W., Suite 550
Washington, DC 20024-2512
Term Expiration: 12/31/98
Phone: (202) 484-3091
FAX: (202) 863-1763
Email: JGARRETT@mathematica-mpr.com

Dr. Malay Ghosh
Department of Statistics
University of Florida
Gainesville, FL 32611-8545
Term Expiration: 12/31/98
Phone: (352) 392-1941 (ext. 232)
FAX: (352) 392-5175
Email: ghoshm@stat.ufl.edu

Dr. F. Thomas Juster
Professor, Institute for Social Research
University of Michigan - Ann Arbor
P.O. Box 1248
Ann Arbor, MI 48106-1248
Term Expiration: 12/31/99
Phone: (313) 764-4207
FAX: (313) 647-1186
Email: ftjuster@umich.edu

Dr. William O'Hare
The Annie E. Casey Foundation
701 St. Paul Street
Baltimore, MD 21202
Term Expiration: 12/31/98
Phone: (410) 223-2949
FAX: (410) 223-2956
Email: billo@aecf.org

Dr. Elizabeth A. Stasny
Department of Statistics
Ohio State University
148D Cockins Hall
Columbus, OH 43210-1247
Term Expiration: 12/31/00
Phone: (614) 292-0784
FAX: (614) 292-2096
Email: eas@stat.mps.ohio-state.edu

Dr. Lynne Stokes
Associate Professor
Department of Management
Science and Information Systems
University of Texas
Austin, TX 78712
Term Expiration: 12/31/00
Phone: (512) 471-5216
FAX: (512) 471-0587
Email: lstokes@mail.utexas.edu

Dr. Roger E. Tourangeau
Gallup Organization
1 Church Street
Rockville, MD 20850
Term Expiration: 12/31/99
Phone: (301) 309-9439
FAX: (301) 309-0635
Email: roger_tourangeau@gallup.com

Committee Liaison: Elizabeth Vacca
Room 3213, Bldg. 4
Phone: (301) 457-4996
FAX: (301) 457-2299
Email: elizabeth.a.vacca@census.gov

APRIL 1998

MEMBERSHIP LIST
CENSUS ADVISORY COMMITTEE OF PROFESSIONAL ASSOCIATIONS
(POPULATION ASSOCIATION OF AMERICA)

Ms. Patricia Becker
APB Associates, Inc.
17321 Telegraph, Suite 204
Detroit, MI 48219
Term Expiration: 4/30/99
Phone: (313) 535-2077
FAX: (313) 535-3556
Email: pbecker@umich.edu

Dr. Linda Jacobsen
Claritas
Suite 400
5375 Mira Sorrento Place
San Diego, CA 92121
Term Expiration: 4/30/98
Phone: (619) 677-9619
FAX: (619) 550-5805
Email: ljacobsen@claritas.com

Dr. Jacob Klerman (Chair)
The Rand Corporation
1700 Main Street
Santa Monica, CA 90406
Term Expiration: 4/30/98
Phone: (310) 393-0411 x6289
FAX: (310) 393-4818
Email: jacob_klerman@rand.org

Dr. Dowell Myers
Associate Professor
School of Planning VKC 351
University of Southern California
Los Angeles, CA 90089-0042
Term Expiration: 4/30/00
Phone: (213) 743-2469
FAX: (213) 743-2476
Email: dowell@usc.edu

Dr. Jeffrey Passel
Program for Research on
Immigration Policy
The Urban Institute
2100 M Street, NW
Washington, DC 20037
Term Expiration: 4/30/99
Phone: (202) 857-8678
FAX: (202) 452-1840
Email: jpassel@ui.urban.org

Dr. Ross Stolzenberg
Department of Sociology
University of Chicago
1155 East 60th Street
Chicago, IL 60637
Term Expiration: 4/30/98
Phone: (847) 835-8451
FAX: (847) 835-2487
Email: stolz@sam.spc.uchicago.edu

Dr. James Trussell
Office of Population Research
21 Prospect Street
Princeton, NJ 08544
Term Expiration: 4/30/98
Phone: (609) 258-4810
FAX: (609) 258-1418
Email: trussell@wws.princeton.edu

Dr. Paul R. Voss
Professor, Dept. of Rural Sociology
University of Wisconsin
311 Agriculture Hall
Madison, WI 53706
Term Expiration: 4/30/00
Phone: (608) 262-9526
FAX: (608) 262-6022
Email: voss@ssc.wisc.edu

Dr. Franklin D. Wilson
Center for Demography and Ecology
University of Wisconsin
1180 Observatory Drive, Room 4412
Madison, WI 53706
Term Expiration: 4/30/99
Phone: (608) 262-2182
FAX: (608) 262-8400
Email: wilson@ssc.wisc.edu

Committee Liaison: Mr. Larry Long
Room 302, Iverson Mall
Phone: (301) 763-8330
FAX: (301) 763-8412
Email: llong@census.gov

APPENDIX E

List of Background Documents

(Asterisk [*] indicates material distributed at meeting)

Agenda for the April 23-24, 1998, Meeting of the Census Advisory Committee of Professional Associations. April 20, 1998. 3 pp.

*Advance Program. [draft] April 17, 1998. 13 pp.

An Overview of Microdata Research using the NSF/Census Survey of Research and Development at the Center for Economic Studies. n.d. 2 pp.

*Balancing and Ratio Editing with the New Speer System. n.d. 6 pp.

*Benchmarking Your R&D: Results from IRI/CIMS Annual R&D Survey for FY '96. n.d. 30 pp.

CENSUS—The Census Bureau More Than Just Numbers. [brochure] March 1998. 14 pp.

*Census 2000 Dress Rehearsal Evaluation Post-Wave Questionnaire. April 1998. 19 pp.

The Census Bureau's Plans for Poverty Measurement Research. April 23, 1998. 15 pp.

*Census 2000 Bulletin, Vol. 2 - No. 24. April 16, 1998. 3 pp.

*Comment—The U.S. Bureau of the Census Corporate Marketing Program. April 23-24, 1998. 3 pp.

*Ron Cooper--comments for CAC meeting, 4/23/98. 4 pp.

*Data Brief—National Science Foundation. December 16, 1997. 4 pp.

*Developing Analytic Programming Capability to Empower the Survey Organization. January 17, 1998. 12 pp.

*Differences in Reported R&D Data on the NSF/Census RD-1 Form and the SEC 10-K Form: Micro-data Investigation. n.d. 27 pp.

*General Edit Imputation Research. n.d. 2 pp.

*Guidelines for Implementing the New Standards. [transparencies] n.d. 9 pp.

How Do We Evaluate the Marketing Strategy in the Dress Rehearsal and Census 2000? April 23-24, 1998. 13 pp.

How Can the Census Bureau Get Consistent and Useful Feedback from its Customers? April 23-24, 1998. 4 pp.

How Should We Promote Confidentiality in the Decennial Census? April 23, 1998. 5 pp., attachment.

How Will The OMB's Preliminary Guidance on Tabulation of Race and Ethnicity Data be Implemented in Dress Rehearsal Tabulations? [transparencies] n.d. 5 pp.

*Key Findings of Corporate Marketing Plan Focus Groups. n.d. 1 p.

*Making Sense of Census 2000. [flyer, student take-home guide] n.d. 4 pp.

Membership List Census Advisory Committee of Professional Associations (American Economic Association). March 1998. 1 p.

Membership List Census Advisory Committee of Professional Associations (American Marketing Association). March 1998. 1 p.

Membership List Census Advisory Committee of Professional Associations (American Statistical Association). March 1998. 1 p.

Membership List Census Advisory Committee of Professional Associations (Population Association of American). March 1998. 1 p.

Multivariate Item Imputation for the 2000 Census Short Form. n.d. 8 pp.

- Balancing and Ratio Editing with the New Speer System. n.d. 6 pp.
- Set-Covering and Editing Discrete Data. n.d. 6 pp.

*Official U.S. Import Statistics on CD-ROM U.S.—Exports History. [product package]

*Official U.S. Import Statistics on CD-ROM U.S.—Imports History. [product package]

Overview of the Census 2000 Dress Rehearsal Evaluation Program. [draft] March 10, 1998. 4 pp., appendixes.

Pricing Products for the Internet: DADS and Beyond. April 23-24, 1998. 5 pp.

*Questions and Issues to be Addressed at the Census AEA Advisory Committee Mini-conference on Technology and Innovation Statistics. n.d. 8 pp.

*R&D: Foundation for Innovation. March-April 1998. 7 pp.

*R&D and Innovation Statistics Census AEA Advisory Committee. April 23, 1998. 5 pp.

Sampling and Estimation in Census 2000 and the Dress Rehearsal. n.d. 18 pp.

Semiannual Summary (Highlights of Developments over the Past 6 Months). April 1998. 25 pp.

*Set-Covering and Editing Discrete Data. n.d. 6 pp.

The Survey of Industrial Research and Development Overview and Recent Developments. April 23, 1998. 10 pp.

- Form RD-1(L)—Survey of Industrial Research and Development During 1997. November 12, 1997. 3 pp.
- Form RD-1L(I)—Instructions for Survey of Industrial Research and Development During 1997, Form RD-1L. November 17, 1997. 11 pp.
- Form RD-1A—Survey of Industrial Research and Development During 1997. October 30, 1997. 4 pp.
- Form RD-1A(I)—Instructions for Survey of Industrial Research and Development During 1997, Form RD-1A. November 12, 1997. 6 pp.

*U.S. Bureau of the Census Center for Economic Studies Annual Report. March 1998. 25 pp.

*U.S. Census Bureau Focus Groups. January 16, 1998. 47 pp.

*U.S. Census Bureau the Official Statistics—CD_ROM Products for the U.S. Census Bureau. [booklet] Fall 1997. 19 pp.

*U.S. Census Bureau the Official Statistics—product profile, LandView® III. [booklet] March 1998. 7 pp.

*United States General Accounting Office—Report to the Committee on Governmental Affairs, U.S. Senate—2000 CENSUS Preparations for Dress Rehearsal Leave Many Unanswered Questions. March 1998. 52 pp.

*U.S. International Trade Data Products Catalog. [brochure] n.d. 20 pp.

The U.S. Bureau of the Census Corporate Marketing Program. April 23-24, 1998. 16 pp.